

Robust Positioning in the Presence of Multipath and NLOS GNSS Signals

Gary A. McGraw, Collins Aerospace, U.S.

Paul D. Groves, University College London, U.K.

Benjamin W. Ashman, National Aeronautics and Space Administration, U.S.

1. Introduction

GNSS signals can be blocked and reflected by nearby objects, such as buildings, walls, and vehicles. They can also be reflected by the ground and by water. These effects are the dominant source of GNSS positioning errors in dense urban environments, though they can have an impact almost anywhere. Non-line-of-sight (NLOS) reception occurs when the direct path from the transmitter to the receiver is blocked and signals are received only via a reflected path. Multipath interference occurs, as the name suggests, when a signal is received via multiple paths. This can be via the direct path and one or more reflected paths, or it can be via multiple reflected paths. Figure 1 illustrates this. Within the GNSS community, it is commonplace to classify NLOS reception as multipath. However, the two effects are not the same; their error characteristics are quite different. As a reflected path is always longer than the direct path, NLOS reception always results in a positive ranging error that is independent of the signal and receiver design. By contrast, the coherent nature of multipath interference can produce both positive and negative ranging errors and these vary with the signal and receiver designs [1].

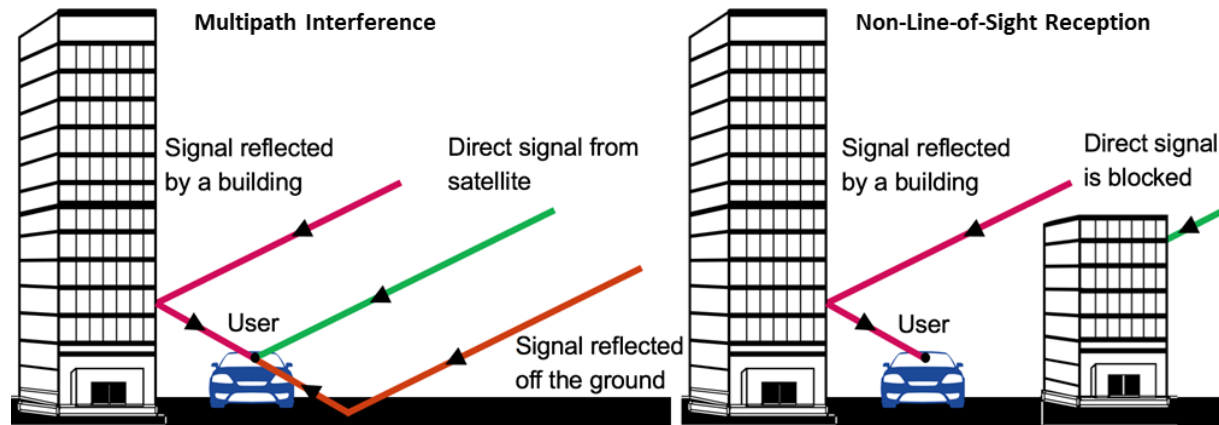


Fig. 1: Multipath interference and NLOS reception.

As their error characteristics are different, NLOS and multipath interference typically require different mitigation techniques, though some techniques are applicable to both. Antenna design and advanced receiver signal processing techniques can substantially reduce multipath errors. Unless an antenna array is used, NLOS reception has to be detected using the receiver's ranging and carrier-power-to-noise-density ratio (C/N_0) measurements and mitigated within the positioning algorithm. Some NLOS mitigation techniques can also be used to combat severe multipath interference. Multipath interference, but not NLOS reception, can also be mitigated by comparing or combining code and carrier measurements, comparing ranging and C/N_0 measurements from signals on different frequencies, and analyzing the time evolution of the ranging and C/N_0 measurements.

Section 2 describes the characteristics of reflected and diffracted signals and how they produce NLOS and multipath errors. Section 3 describes how multipath errors can be reduced using advanced receiver design and signal-processing techniques, including antenna design considerations, correlation signal processing and adaptive antenna array processing. Section 4 covers carrier smoothing of code measurements, which is a technique for mitigating both noise and multipath. Section 5 describes real-time navigation-processor-based NLOS and multipath mitigation techniques, including C/N_0 -based detection and weighting, outlier detection, and aiding from other sensors. Section 6 then describes multipath mitigation techniques for post-processed high-precision positioning that work by analyzing time series of GNSS measurement data. Finally, Section 7 describes three-dimensional-mapping-aided (3DMA) GNSS. This improves real-time positioning in dense urban environments by using 3D mapping to predict which signals are NLOS at which locations. This can be used to enhance conventional ranging-based positioning and to implement shadow matching, a complementary GNSS positioning technique that determines position by comparing predicted and measured C/N_0 from several satellites.

2. Characteristics of Reflected Signals, NLOS Reception, and Multipath Errors

2.1 Multipath/NLOS Signal Characteristics

For land applications, most GNSS signal reflections occur within the surrounding environment, such as the ground, buildings, vehicles, or trees. For air, sea, and space applications, reflections off the host-vehicle body are more common. Low-elevation-angle signals are more likely than high-elevation-angle signals to be received via reflections by vertical surfaces. Where the reflecting surface is rough compared to the signal wavelength ($\sim 0.2\text{m}$ for GNSS), scattering occurs, resulting in weak signals being reflected in many different directions. Within a GNSS receiver, scattered signals typically manifest as additional noise. Where the reflecting surface is smooth and of sufficient size, specular reflection occurs whereby a strong signal is reflected at an angle equal and opposite to the angle of incidence of the direct signal at the reflecting surface.

GNSS signals are right-hand circularly polarized (RHCP). Specular reflection from a surface at normal incidence results in a left-hand circularly polarized (LHCP) reflected signal. At other angles of incidence, the polarization is mixed. As the angle of incidence increases, the LHCP component of the reflected signal decreases and the RHCP component increases. At Brewster's angle the two components are equal, while at larger angles, the RHCP component dominates. The value of Brewster's angle depends on the frequency of the signal and the properties of the reflecting surface. For L-Band GNSS signal ($\sim 1.5\text{ GHz}$), it is around 89° for metallic surfaces, 85° for sea water and around 70° for soil [2][3].

Reflected signals are always delayed with respect to direct signals and have a lower amplitude (unless the direct signals are blocked or attenuated). In cases of NLOS reception, the ranging errors are potentially unbounded and always positive. Thus, errors of several kilometers occasionally occur when a signal is reflected by a distant tall building. The range-rate errors can result in the user's apparent direction of motion being reflected in the object reflecting the signal. Consequently, reflectors perpendicular to the direction of travel can produce much larger velocity errors than those parallel to the trajectory. Also, for a reflector close to a moving user antenna, the pseudorange error may be small, but the range-rate error large.

The signal path between satellite and user is not a simple ray, but is instead determined by Fresnel zones. The first Fresnel zone is defined as the region about a signal's propagation path where the phase difference in path length is less than half a cycle and is the radius of the effective signal footprint where the signal

interacts with an object in its path. For reflection geometries where the transmitter is quite distant from both the receive antenna and the reflecting surface, as is the case with GNSS signals, the first Fresnel zone can be approximated as $\sqrt{r\lambda_L}$, where r is the distance of the object from the user antenna and λ_L is the carrier wavelength [5]. Irregularities in the object on this scale will therefore affect the properties of the reflected or diffracted signal. Where a signal is partially blocked by an obstacle, diffraction can occur. The part of the signal interacting with the object interferes with the part passing the object by re-radiating energy in different directions and with different phase shifts, bending the path of the signal and producing alternating bands of constructive and destructive interference. The difference between the maximum and minimum levels of this diffraction pattern reduces as the angle through which the signal bends to pass around the obstacle increases, with useable GNSS signals receivable at deflections of up to 5° [4]. Diffracted signals are also delayed, but typically only by decimeters, as the strongest scattering occurs along paths that only deviate slightly from the direct path.

At the receiver, each reflected signal may be described by a relative amplitude, α_i , range lag, Δ_i , and carrier phase offset, ϕ_i , with respect to the direct signal (or the strongest signal if no direct signal is received). There is also a carrier frequency offset, δf_i , which is larger when the user is moving with respect to the reflecting surface [5]. The relative amplitude of the i^{th} reflected signal is given by

$$\alpha_i = \sqrt{\frac{G_i R_i k_i}{G_0 R_0 k_0}} \quad (1)$$

where G_i and G_0 are the antenna gains in the directions of the i^{th} and strongest signals, respectively, R_i and R_0 are the reflection coefficients, and k_i and k_0 are the path attenuation coefficients. The path attenuation coefficients are approximately 1 unless one of the signals has been attenuated by passing through partially absorbing materials, such as foliage. When the strongest signal is the direct signal, $R_0 = 1$. For reflected signals, the reflection coefficient depends on the properties of the reflecting surface. Calm water, metal, and metallized glass can produce particularly strong specular reflections with reflection coefficients of 0.5-0.7. Brick, stone and concrete typically have lower reflection coefficients. Note that rainwater enhances the reflectivity of other surfaces, including the ground, walls and foliage.

The antenna gain is a function of the signal's angle of incidence at the antenna and of its polarization. In practice, there can also be variation in the gain with azimuth. A typical GNSS antenna has a gain of 1.6-2.5 (2-4 dB) for signals at normal incidence (i.e., at the antenna zenith). This drops as the angle of incidence increases and is generally less than 1 for angles of incidence greater than 75° . For a horizontally mounted antenna, a 75° incidence angle corresponds to a satellite signal at a 15° elevation angle. Placing the antenna on a ground plane, significantly attenuates signals from below the plane of the antenna, reducing the impact of ground reflections. Fixed site applications often use special designs, such as choke ring antennas that attenuate low-elevation (i.e., high-incidence) signals.

GNSS antennas are often designed to have higher gain for RHCP signals. At normal incidence to an RHCP planar antenna, the gain for LHCP signals is about a factor of 10 (10 dB) less than the RHCP gain. The polarization discrimination drops as the angle of incidence increases such that the LHCP and RHCP gains are the same for signals incident parallel to the plane of the antenna. An RHCP GNSS antenna is effectively linearly polarized for signals from the side.

Mobile phone antennas, conversely, are usually linearly polarized to minimize size. These antennas offer no polarization discrimination at all, so smartphone receivers are more susceptible to both multipath interference and NLOS reception. Furthermore, signals with lines-of-sight along the axis of a dipole antenna are significantly attenuated, which can sometimes result in a reflected signal being stronger than the direct counterpart.

The path delay is the additional distance travelled by a reflected signal compared to the direct path from the satellite to the receiver. Figure 2 shows the signal paths of direct, building-reflected, and ground-reflected signals. For the building-reflected signal, the range lag is $\Delta_i = a + b$, while for the ground-reflected signal it is $\Delta_i = e - g$. It is also useful to define the path delay in code chips, $\delta_i = \Delta_i / \lambda_c$, where λ_c is the code chip length. The phase offset is given by [6]

$$\phi_i = \left(\frac{2\pi\Delta_i}{\lambda_L} + \phi_{Ri} \right) \text{MOD} 2\pi \quad (2)$$

where λ_L is the carrier wavelength, the MOD operator gives the remainder from division by 2π , and ϕ_{Ri} is the phase shift on reflection, which is π radians for a totally flat specular reflector at an angle of incidence less than Brewster's angle. The frequency offset from the direct signal carrier frequency is [6]

$$\delta f_i = \frac{1}{2\pi} \frac{\partial \phi_i}{\partial t} = \frac{\partial}{\partial t} \left(\frac{\Delta_i}{\lambda_L} + \frac{\phi_{Ri}}{2\pi} \right). \quad (3)$$

In the case of a building-reflected signal, the range lag is $\Delta_i = 2d \cos \theta$ for satellite elevation θ and distance from the reflecting wall d . The resulting multipath fading frequency is

$$\delta f_i = \left(\frac{2}{\lambda_L} \cos \theta \right) \frac{\partial d}{\partial t} - \left(\frac{2d}{\lambda_L} \sin \theta \right) \frac{\partial \theta}{\partial t}. \quad (4)$$

In the case of a ground-reflected signal, the range lag is $\Delta_i = 2h \sin \theta$ for antenna height h and the resulting multipath fading frequency is

$$\delta f_i = \left(\frac{2}{\lambda_L} \sin \theta \right) \frac{\partial h}{\partial t} + \left(\frac{2h}{\lambda_L} \cos \theta \right) \frac{\partial \theta}{\partial t}. \quad (5)$$

Thus, the multipath fading frequency can be seen to be a function of the distance from the reflecting surface (i.e., d or h), along with the rate of change of the distance. In a case experiencing both types of multipath, as in Fig. 2, the relative influence of d and h on the composite fading frequency is inversely proportional to their relative magnitudes—for a more extensive treatment, see [7].

Although this is a simplified scenario, some useful conclusions can be drawn. For a stationary antenna with ground reflections, for example, if h is 1 meter and $\partial \theta / \partial t$ is approximately $180^\circ / 6$ hours (~ 0.15 mrad/s), then the multipath fading frequency at low elevation angles is about 1.5 mHz. However, if the antenna is moving perpendicular to a reflecting wall at 1 m/s, such as a car backing out of a parking space, $2(\partial d / \partial t) / \lambda_L$ is on the order of 10 Hz. At a rate of perpendicular motion of only 10 m/s, δf_i measures tens of Hz. Fading frequencies much greater than 10 Hz would typically exceed the carrier tracking loop bandwidth, appearing more as noise than an error source correlated with the geometry of the antenna environment. For vehicles at highway speeds, multipath from stationary objects is usually insignificant, but

in stop and go traffic such multipath may be a problem. Multipath is much more of a concern when the reflector is stationary with respect to the antenna, as in the case of fixed terrestrial installations or on spacecraft or aircraft when structures on the vehicle are present near the antenna.

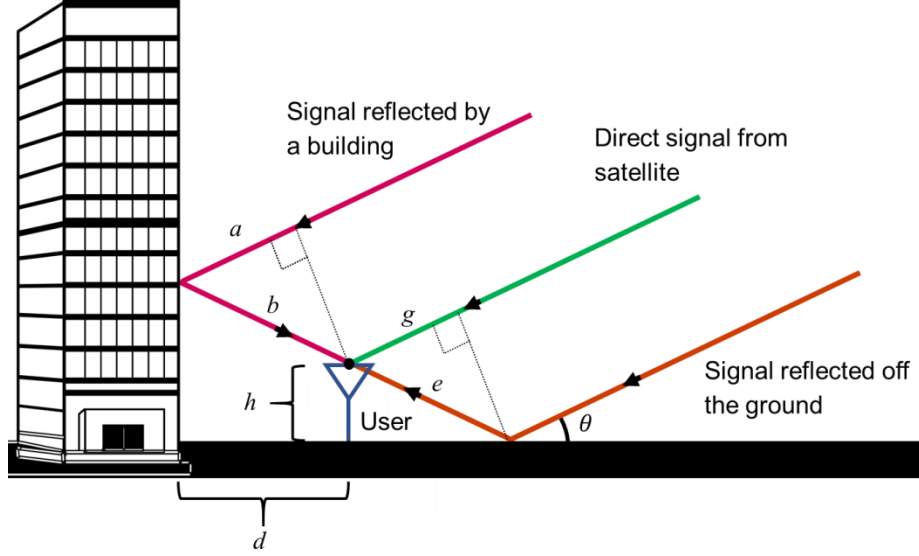


Fig. 2: Direct, building-reflected and ground-reflected signal paths in a multipath interference scenario.

2.2 Receiver Signal Modeling

The total received GNSS signal after the antenna can be written as [6],[8],[9]:

$$r(t) = \sqrt{2P} \sum_{i=0}^n \left\{ \alpha_i C(t - t_0 - \Delta_i/c) D(t - t_0 - \Delta_i/c) \times \exp\left(j \left[2\pi (f_L + f_D + \delta f_i)(t - t_0 - \Delta_i/c) + \phi_i \right] \right) \right\} + w(t) \quad (6)$$

where P is the signal power, n is the number of reflected signals, $C(\cdot)$ is the pseudorandom noise (PRN) spreading code, $D(\cdot)$ is the downlink data, t_0 is the propagation delay for the direct signal, c is the speed of light, f_L is the carrier frequency, f_D is the Doppler shift, ϕ_0 is the carrier phase of the direct signal component, and $w(t)$ is bandlimited white Gaussian noise (WGN). In a typical receiver there is usually a mixing operation to an intermediate frequency (IF), but we will assume a direct conversion to baseband for modeling purposes. In addition, from this point on, the navigation data in (6) is omitted since it generally has little impact on multipath error.

In an idealized basic GNSS receiver, illustrated in Fig. 3, incoming signals are converted to baseband with the in-phase and quadrature numerically-controlled oscillator signals (NCO), $s_{NCO}(t)$, and then correlated with early, prompt and late replica codes, s_E , s_P , and s_L , respectively:

$$s_{NCO}(t) = \exp\left(-j \left[2\pi (f_L + \hat{f}_D) t + \hat{\phi} \right] \right) \quad (7)$$

$$s_P(t) = C(t - \hat{t}_0), \quad s_E(t) = C(t + dT_C - \hat{t}_0), \quad s_L(t) = C(t - dT_C - \hat{t}_0)$$

where d is the early-prompt and prompt-late correlator spacing in units of chips, and \hat{t}_0 , \hat{f}_D , and $\hat{\phi}$ are estimated quantities. The correlation process is assumed to include an automatic gain control (AGC) normalization by noise power and the coherent pre-detection integration interval, T_{PDI} .

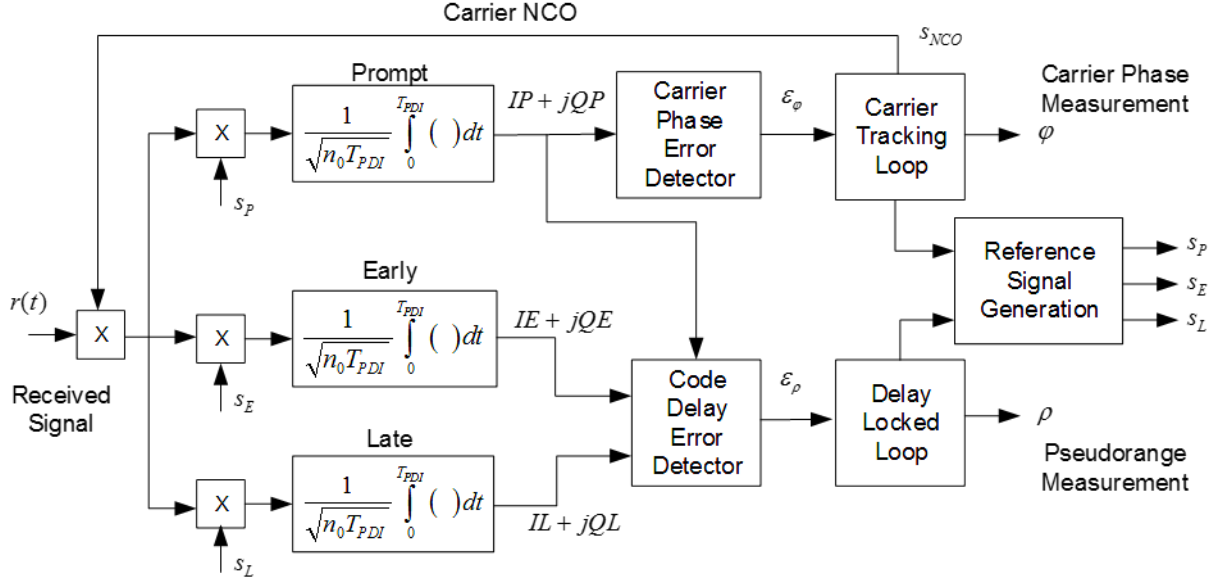


Fig. 3: Receiver signal processing block diagram.

In the presence of multipath interference where the composite signal is given by (6), the accumulated correlator outputs become [6],[8],[9]:

$$\begin{aligned}
 IE &= \sqrt{2(c/n_0)T_{PDI}} \sum_{i=0}^n \left[\alpha_i R(\tau - \delta_i T_C + dT_C) \text{sinc}(\pi(\delta f + \delta f_i)T_{PDI}) \cos(\delta\varphi + \phi_i) \right] + w_{IE} \\
 IP &= \sqrt{2(c/n_0)T_{PDI}} \sum_{i=0}^n \left[\alpha_i R(\tau - \delta_i T_C) \text{sinc}(\pi(\delta f + \delta f_i)T_{PDI}) \cos(\delta\varphi + \phi_i) \right] + w_{IP} \\
 IL &= \sqrt{2(c/n_0)T_{PDI}} \sum_{i=0}^n \left[\alpha_i R(\tau - \delta_i T_C - dT_C) \text{sinc}(\pi(\delta f + \delta f_i)T_{PDI}) \cos(\delta\varphi + \phi_i) \right] + w_{IL} \\
 QE &= \sqrt{2(c/n_0)T_{PDI}} \sum_{i=0}^n \left[\alpha_i R(\tau - \delta_i T_C + dT_C) \text{sinc}(\pi(\delta f + \delta f_i)T_{PDI}) \sin(\delta\varphi + \phi_i) \right] + w_{QE} \\
 QP &= \sqrt{2(c/n_0)T_{PDI}} \sum_{i=0}^n \left[\alpha_i R(\tau - \delta_i T_C) \text{sinc}(\pi(\delta f + \delta f_i)T_{PDI}) \sin(\delta\varphi + \phi_i) \right] + w_{QP} \\
 QL &= \sqrt{2(c/n_0)T_{PDI}} \sum_{i=0}^n \left[\alpha_i R(\tau - \delta_i T_C - dT_C) \text{sinc}(\pi(\delta f + \delta f_i)T_{PDI}) \sin(\delta\varphi + \phi_i) \right] + w_{QL}
 \end{aligned} \tag{8}$$

where I and Q denote in-phase and quadra-phase, E , P and L denote early, prompt and late, c/n_0 is the carrier-power-to-noise-density ratio (non-decibel form), $R(\cdot)$ is the autocorrelation function of the PRN sequence, $\tau = \hat{t}_0 - t_0$ is the code tracking error in units of seconds, δf is the carrier frequency tracking error, $\delta\varphi$ is the carrier phase tracking error, and without loss of generality it is assumed that the AGC normalizes the I/Q noise terms $w_{IE}, w_{IP}, w_{IL}, w_{QE}, w_{QP}$ and w_{QL} to have zero mean and unit variance. Note that

$$\text{sinc}(\theta) = \begin{cases} \sin \theta / \theta, & \theta \neq 0 \\ 1, & \theta = 0 \end{cases} \quad (9)$$

See Chapters 14 and 15 of this volume, and [6],[8],[9],[11] or [12] for further details of GNSS receiver signal processing. For an ideal PRN code, the autocorrelation function is

$$R(\tau) = E\{C(t)C(t-\tau)\} = \begin{cases} 1-|\tau/T_C|, & |\tau| < T_C \\ 0, & |\tau| \geq T_C \end{cases} \quad (10)$$

In subsequent sections, we will also make use of the Early-Minus-Late (EML) Delay Lock Detector (DLD) function:

$$D_{EML}(\tau) = \frac{1}{2}[R(\tau + dT_C) - R(\tau - dT_C)] \quad (11)$$

The factor of half yields unity slope near the origin for the infinite bandwidth case, giving an accurate measurement of the tracking error.

2.3 Code Multipath Error

The effect of multipath on the code pseudorange measurements is principally a function of how multipath manifests itself in the code delay error detector in Fig. 3. Two common delay error detectors, that are power- and phase-insensitive approximations of (11), include the Early-Power Minus Late-Power discriminator:

$$\varepsilon_{ELP} = \frac{(IE^2 + QE^2) - (IL^2 + QL^2)}{2(IP^2 + QP^2)} \quad (12)$$

and the Dot-Product detector:

$$\varepsilon_D = \frac{ID_{EML} \cdot IP - QD_{EML} \cdot QP}{IP^2 + QP^2}, \quad (13)$$

$$ID_{EML} = (IE - IL) / 2, \quad QD_{EML} = (QE - QL) / 2$$

Fig. 4(a) shows the direct-signal, reflected-signal and combined code autocorrelation functions for an ideal biphasic shift key (BPSK) GNSS signal subject to a single interfering signal with a relative amplitude, $\alpha_m = 0.4$, a path delay in code chips, $\delta_m = 0.125$ code chips and phase offsets, ϕ_m , of 0 and 180°. The correlator spacing is $d = 0.25$ chip. The plots are done using a normalized time offset $\bar{\tau} = \tau / T_C$ to make the results independent of PRN chipping rate. The effect of pre-correlation band-limiting on the shape of the correlation function is also neglected in this diagram. It can be seen that the multipath interference distorts the shape of the correlation function. A conventional code-tracking loop works by adjusting the phase of the receiver generated code so that the signal powers in the early and late correlation channels are equal. Therefore, if the code autocorrelation function is asymmetric due to multipath interference, there will be a code-tracking error and, consequently, an error in the resulting pseudorange measurement. Fig. 4(b) shows the EML detector functions as in (11) corresponding to the case in Fig. 4(a). It can be seen that the zero crossings for the direct-plus-multipath cases are biased from the nominal case.

Assuming that the multipath fading frequency is within the Delay-Locked Loop (DLL) code tracking bandwidth, determining the code multipath error involves computing where the delay error detector $\varepsilon_\rho = 0$. In general this is best done with numerical methods, i.e., computing (12) or (13) with the direct and

multipath signal components. However, an analytical solution can be readily obtained for the case of infinite bandwidth and a single interfering signal. For the dot product detector (13) this can be done by solving:

$$D_{EML}(\bar{\tau}) - \alpha_1 D_{EML}(\bar{\tau} - \delta_1) \cos \phi_1 = 0 \quad (14)$$

The pseudorange error due to multipath is obtained from the solution of (14) as $\delta \rho_M = \bar{\tau} T_C$. For example, for a short multipath delay such that $|\bar{\tau} - \delta_1| < d$, then (14) yields

$$\delta \rho_M = \frac{-\alpha_1 \cos \phi_1}{1 - \alpha_1 \cos \phi_1} \delta_1 T_C = \frac{-\alpha_1 \cos \phi_1}{1 - \alpha_1 \cos \phi_1} \Delta_1 \quad (15)$$

In this case the pseudorange error can be seen to be proportional to the multipath delay and amplitude. Upper and lower error bounds for the dot product detector can be obtained by solving (14) for $\phi_1 = 0$ or π . The computed bounds for BPSK code modulation are indicated in Fig. 5. The plot also shows the errors for the specific correlator spacing in Fig. 4(b). Also shown in Fig. 5 is a plot of the multipath bias, which was computed numerically by averaging the result of (14) over $-\pi < \phi_1 \leq \pi$ for each value of δ_1 .

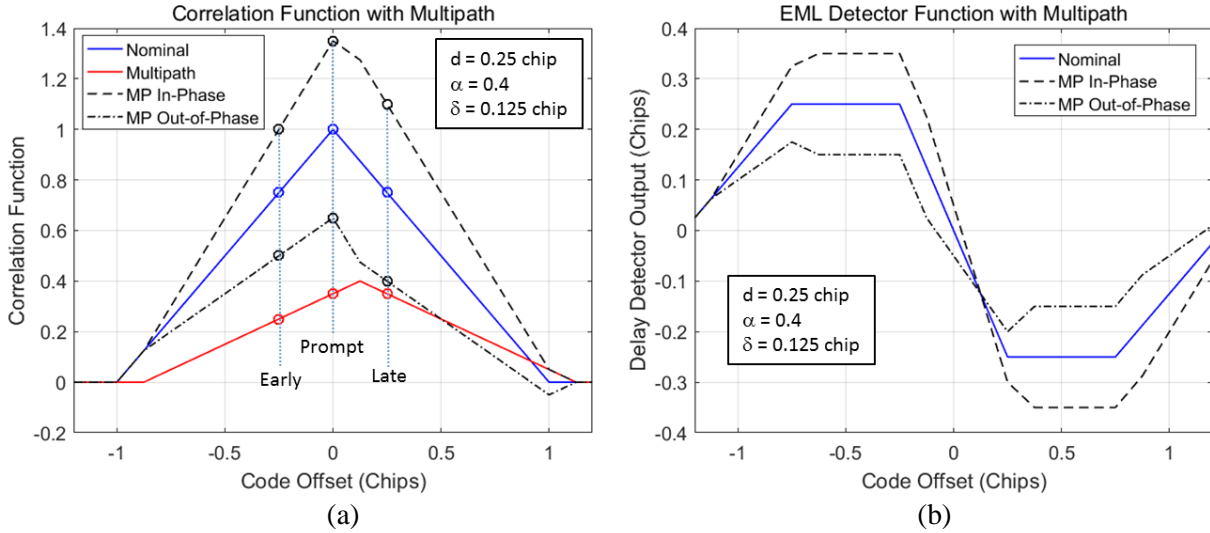


Fig. 4: Correlation and delay-detector functions of a BPSK GNSS signal subject to constructive and destructive multipath interference ($d = 0.25$ chips; $\alpha_1 = 0.4$; $\delta_1 = 0.125$ chips; pre-correlation band-limiting is neglected).

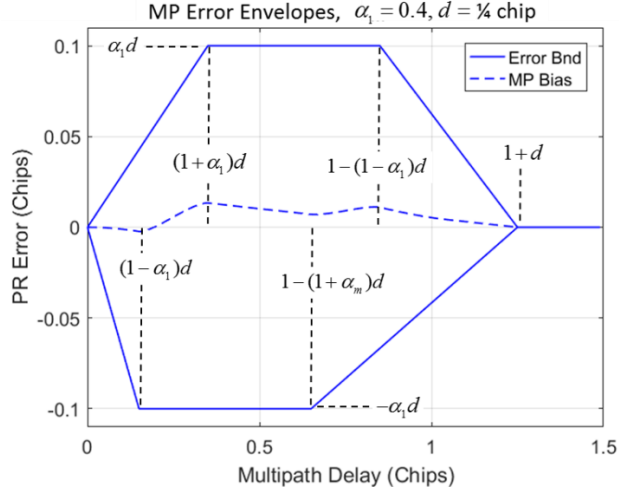


Fig. 5: Multipath error envelopes and multipath bias for a BPSK signal and EML tracking for a single multipath signal. Pre-correlation band-limiting is neglected.

As seen in Fig. 5, the maximum absolute multipath error is $\alpha_1 d$, so a narrower correlator spacing often leads to a smaller tracking error. This is one of the benefits of narrow correlator receiver designs [9]. For an ideal PRN code, the multipath error is zero if $\delta_1 > 1 + d$, however for codes with non-negligible autocorrelation side peaks a small residual error is possible [5],[95]. The code multipath error at any given instance will be between the bounds shown in Fig. 5 and will oscillate at a frequency determined by the rate of change in the direct and reflected signal geometry as discussed above. As the multipath fading frequency increases beyond the carrier tracking loop capability, the code multipath error will converge to the bias level shown in Fig. 5.

Higher chipping-rate signals are less susceptible to multipath interference than low chipping-rate signals as the range lag, Δ , must be smaller for the reflected signal to affect the correlation peak. Figure 6(a) compares the code multipath error envelopes for three different modulation types: BPSK(1), BPSK(10) and binary offset carrier BOC(1,1), assuming infinite bandwidth. It can be seen that the BPSK(10) signal has much better resistance to longer delay multipath. However in many high accuracy applications, it is the short delay multipath— with path delay less than 0.1 BPSK(1) chip (29.3 m for C/A code)—that is the dominate error source, and all the code types have similar errors.

Real GNSS receivers are affected by pre-correlation filtering, which is necessary to eliminate out of band interference and band-limit the signal spectrum prior to sampling. For high chipping-rate signals, the band-limiting at the transmitter is also significant. Figure 6(b) shows the corresponding simulated error envelopes for the signals in Fig. 6(a) with the code signals filtered by a five-pole 10 MHz low-pass filter (corresponding to a 20 MHz band-pass filter in the receiver front end) and the dot-product detector (13). The peak multipath errors are actually slightly reduced from the infinite bandwidth case, but the broadening of the correlation function extends the error bounds, which can be observed in Fig. 6 for BPSK(10). Band-limiting rounds the correlation function, which reduces the benefit of narrowing the correlator spacing, as this causes the slope of the delay error detector to decrease. Band-limiting also reduces the effectiveness of the receiver signal-processing-based multipath mitigation techniques described in Section 3.2.

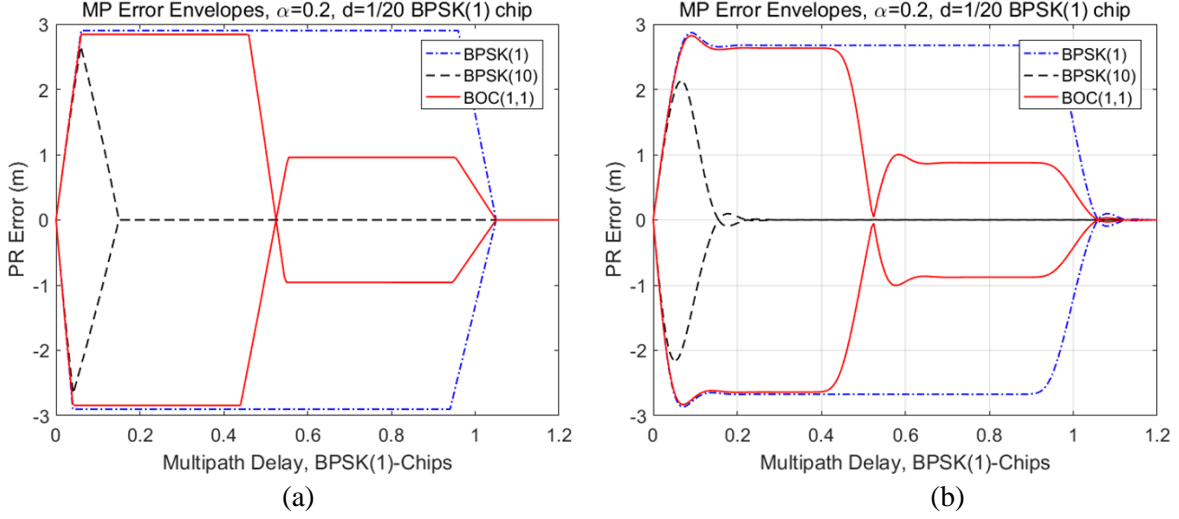


Fig. 6: Code multipath error envelopes for different code types. (a) Infinite bandwidth; (b) Filtered with 10 MHz low-pass filter.

2.4 Carrier Multipath Error

In most GNSS receivers, carrier tracking relies on a phase tracking error measurement from the prompt correlator signal. The carrier tracking loop attempts to drive this tracking error to zero. For a single reflected signal, the carrier phase error can be determined using a simple signal phasor model as depicted in Fig. 7, yielding:

$$\delta\varphi_M = \arctan\left(\frac{\alpha_1 R(\bar{\tau} - \delta_1) \sin \phi_1}{R(\bar{\tau}) + \alpha_1 R(\bar{\tau} - \delta_1) \cos \phi_1}\right) \quad (16)$$

The peak carrier phase error can be derived from (16), which occurs when $\cos \phi_1 = -\alpha_1 R(\bar{\tau} - \delta_1) / R(\bar{\tau})$:

$$\delta\varphi_{Max} = \arctan\left(\frac{\alpha_1 R(\bar{\tau} - \delta_1)}{\sqrt{R(\bar{\tau})^2 - \alpha_1^2 R(\bar{\tau} - \delta_1)^2}}\right) \quad (17)$$

Unlike the code error, the worst-case carrier phase error happens when $\delta_1 = 0$ and for $\alpha_1 < 1$ the carrier phase error does not exceed 90° , corresponding to 4.8 cm at the L1 carrier frequency. Similar to what was done for code multipath, error bounds can be computed from (16). The results for $\alpha_1 = 0.2$ are shown in Fig. 8(a) for the infinite bandwidth case and Fig. 8(b) for a filtered case. As in the case of code multipath, the actual errors will oscillate between these extremes.

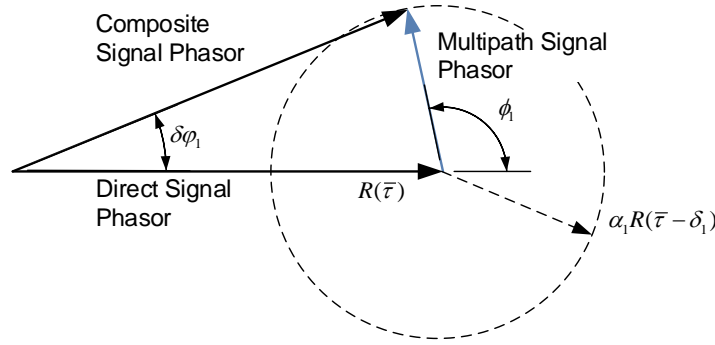


Fig. 7: Depiction of a carrier phase multipath error for a single reflected path. The relative phase of the multipath signal with the direct path dictates the measured carrier phase error.

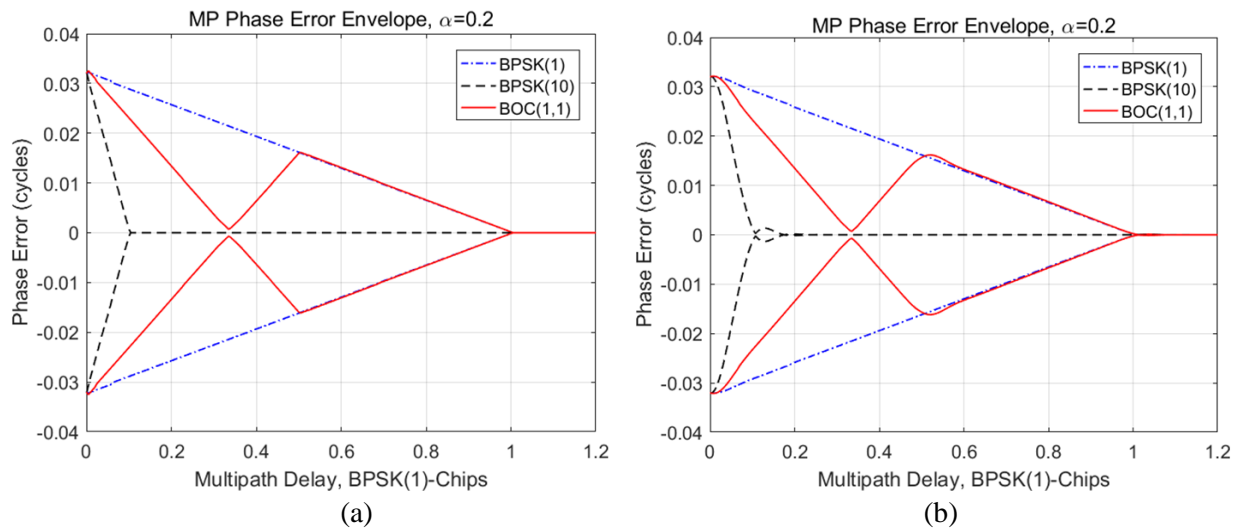


Fig. 8: Carrier multipath error envelopes for different code types. (a) Infinite bandwidth; (b) Filtered with 10 MHz low-pass filter.

3. Receiver-Based Multipath Mitigation

Mitigation of multipath in GNSS signals can be accomplished throughout the receiver signal processing chain, including the antenna. This section covers receiver multipath mitigation techniques that are hosted in the receiver up to the point of code and carrier measurement generation. Note that none of these techniques directly mitigate NLOS reception errors. Measurement-domain mitigation techniques are discussed in Sections 4 and 5.

3.1 Antenna Design Techniques

Attenuating multipath interference through antenna design techniques can be highly effective in many applications. Two general approaches are discussed here: 1) Enhancement of desired-to-undesired signal levels in fixed antenna response characteristics; and 2) Taking advantage of polarization diversity in the antenna to generate measurements of multipath parameters.

3.1.1 Desired/Undesired Signal Component Optimization

Attenuating multipath interference prior to entering receiver signal processing is highly desirable when this is possible. Primary examples where multipath mitigation drives antenna design are fixed site, survey,

aircraft and other vehicular applications where the multipath generally arrives below the receiver mask angle. Figure 9 illustrates an idealized antenna response wherein gain in the direction of the satellite is enhanced and gain in the direction of multipath is attenuated. As discussed in detail in Chapter 26, enhancement of the desired/undesired signal ratio can be accomplished by a variety of design elements, including ground planes, choke ring assemblies and spiral antenna elements.

These multipath mitigating antenna design features tend to increase the size of the antenna, limiting their application. An extreme example is the Integrated Multipath Limiting Array (IMLA) for the Ground Based Augmentation System (GBAS) [16] which is a special-purpose fixed-site design that approximates the ideal fixed antenna pattern by combining two separate antenna structures: 1) The MLA, which consists of a vertical array of dipole elements in a manner that creates a sharp cut off in the gain pattern below five degrees and provides high gain up to a 35 degree elevation angle; 2) A High Zenith Antenna (HZA) which provides for satellite tracking above 35 degrees and more than 30 dB of attenuation of ground multipath reflections. The IMLA is more than a meter high and is typically installed an additional meter or more above the ground.

For many applications, such as handheld devices, the antenna must be small and the device orientation can be almost arbitrary, so it is essentially impossible to include multipath limiting features, hence other multipath mitigation techniques must be pursued.

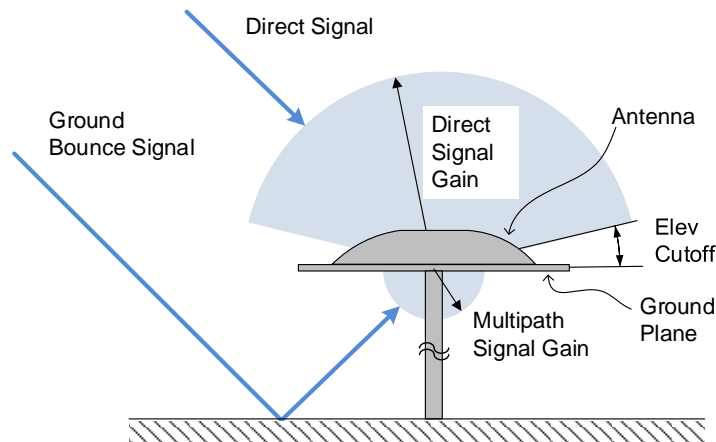


Fig. 9: Illustration of an idealized antenna pattern where the direct signal gain is enhanced and the multipath interference is attenuated.

3.1.2 Polarization Diversity Reception

As discussed in Section 1, transmitted GNSS signals are RHCP. In general, this circular polarization is lost upon reflection, but the resulting elliptically-polarized signal is mostly LHCP if the grazing angle is greater than Brewster's angle (i.e., the angle of incidence is sufficiently small). Higher performance GNSS antennas are typically tuned to receive RHCP signals, thereby already providing some resistance to multipath, although RHCP selectivity falls off at low elevation angles. Further improvement can be achieved, however, by employing an LHCP antenna to preferentially receive reflected signals and incorporate this additional information. Multipath mitigation techniques in this class can be categorized as measurement weighting, range domain correction, tracking domain correction, and adaptive antenna array processing. Measurement weighting and antenna array techniques are covered in Sections 5.1 and 3.4, respectively. Tracking domain correction has been shown in simulation [17][18], but relies on prohibitively detailed knowledge of the

antenna gain patterns and the direction of the reflected signal arrival relative to the antennas [19]. Range domain correction is discussed briefly here.

Pseudorange, carrier phase and C/N_0 measurements can be made from both polarizations. Estimation of multipath using dual polarization techniques relies on differences in the behavior of direct and reflected signals under the two antenna polarizations. Under certain conditions, the multipath parameters α_i , Δ_i and ϕ_i can be estimated from the RHCP and LHCP measurements and then used to estimate the code and carrier errors, $\delta\rho_M$ and $\delta\varphi_M$. These error estimates can in turn be used to correct range measurements supplied to the navigation processor.

Pseudorange and carrier phase errors are the result of the relative delay and amplitude of reflected signals, as indicated in Equations (14) and (16), respectively. However, these errors also depend on the direct/reflected signal mix produced by the polarization of the reflected signal and the cross-polarization discrimination of the antenna. Assuming good polarization reversal at reflection (i.e., reflected signals are LHCP), the pseudorange error due to a single short-delay reflected signal in (15) can be expanded

$$\delta\rho_M = \frac{-\sqrt{\frac{G_L}{G_R}}\bar{\alpha}_1 \cos \phi_1}{1 - \sqrt{\frac{G_L}{G_R}}\bar{\alpha}_1 \cos \phi_1} \Delta_1 \quad (18)$$

where the cross-polarization discrimination of the antenna — the ratio of the LHCP and RHCP gains, G_L and G_R , respectively [20] — has been separated from the relative amplitude defined in (1), such that $\bar{\alpha}_i = \sqrt{G_R/G_L}\alpha_i$. For a very good LHCP antenna, G_L/G_R is large and in the limit $\delta\rho_M$ approaches Δ_1 , the multipath relative delay. In this case of good isolation, the difference between the measured LHCP and RHCP pseudoranges is approximately the geometric additional path length of the reflected signal relative to the direct:

$$\begin{aligned} \rho_{LHCP} - \rho_{RHCP} &= (\rho_0 + \Delta_1) - (\rho_0 + \delta\rho_M) = \Delta_1 + \frac{\alpha_1 \cos \phi_1}{1 - \alpha_1 \cos \phi_1} \Delta_1 \\ &\approx \Delta_1, \quad \alpha_1 \ll 1 \end{aligned} \quad (19)$$

For carrier phase, (16) can be expanded in a manner similar to what was done in (18), yielding

$$\begin{aligned} \varphi_{LHCP} - \varphi_{RHCP} &= (\phi_0 + \phi_1) - (\phi_0 + \delta\varphi_M) = \phi_1 - \arctan\left(\frac{\alpha_1 \sin \phi_1}{1 + \alpha_1 \cos \phi_1}\right) \\ &\approx \phi_1, \quad \alpha_1 \ll 1 \end{aligned} \quad (20)$$

For signal-to-noise ratio (SNR) measurements with a single multipath signal component, as illustrated in Fig. 7, the ratio of the LHCP and RHCP SNR is [78]:

$$\begin{aligned} \frac{SNR_{LHCP}}{SNR_{RHCP}} &= \frac{\alpha_1^2}{1 + \alpha_1^2 + 2\alpha_1 \cos \phi_1} \\ &\approx \alpha_1^2, \quad \alpha_1 \ll 1 \end{aligned} \quad (21)$$

The GNSS observables in (19), (20) and (21), have been used as bistatic radar measurements in aircraft altimetry [21][22][23] and relative navigation applications [24][25][26][20].

Isolation of these multipath parameters requires that the cross-polarization discrimination of the antennas renders the reception of oppositely-polarized signals negligible. In the case of interest here, in which we seek to correct multipath-induced errors in LOS signal ranging, this assumption is only approximately true, so the multipath correction is also only approximate. To truly correct the multipath errors not only requires knowledge of the path delay from (19), but also that the multipath phase, ϕ_i , be determined to within about one tenth of a carrier cycle (~ 2 cm), so the effects of antenna separation must be considered, or a single dual-polarized antenna must be used [19]. See Sections 3.3 and 6.2 for further discussion of the estimation of multipath parameters on which range domain corrections rely.

3.2 Correlation Signal Processing Techniques

It was recognized early in the development of GPS that authorized (military) users would obtain superior multipath mitigation with the BPSK(10) P(Y)-code compared to the BPSK(1) C/A-code available to all users, as illustrated in Figures 6 and 8. The widespread adoption of the GPS C/A-code for numerous commercial and civil applications in the 1990s and early 2000s spurred innovations in multipath mitigation using a variety of signal processing methods. In this section, correlator signal processing techniques will be considered; parameter estimation and array processing techniques are discussed in subsequent sections.

The literature abounds with a variety of different techniques, see References [27]-[35] for a sampling of the techniques. The narrow correlator [9] technique was among the earliest improved correlator signal processing methods. This technique was described in Section 2.3, where it was shown that, for sufficient signal bandwidth, reducing the correlator spacing, d , to a small fraction of a code chip results in reduction of peak multipath errors. We concentrate on a generic class of techniques that we call “Signal Gating”, although manufacturers have different names for their specific implementations. The focus is on code multipath mitigation, although carrier phase multipath mitigation has been pursued with these methods as well.

3.2.1 Signal Gating

The concept of signal gating is illustrated in Fig. 10. The top signal represented in Fig. 10 is the incoming BPSK PRN code which is to be tracked using a DLD function that provides a measurement of code tracking error. Instead of differencing early and late correlation functions, as illustrated in Fig. 3, the EML DLD function in (11) can also be generated from rectangular signal gates around transitions of the PRN code, as illustrated by the “Conventional DLD Signal” in Fig. 10. In essence, the Early and Late reference signals s_E and s_L in (7) are replaced by a DLD signal, s_D , that is then correlated with the incoming satellite signal. In signal gating, this DLD signal is manipulated to minimize multipath. To aid in the derivation of DLD functions for the different types of signal gates, we will first re-derive the EML DLD function based on the signal gating concept. This process is illustrated in Fig. 11.

In Fig. 11, two PRN codes epochs, t_k and t_{k+1} are considered. We assume that after epoch t_{k-1} that the PRN code value is +1. Then there are four possible PRN code state transitions that can occur at t_k and t_{k+1} as shown. When these signals are correlated with the EML DLD signal of width $2d_1T_C$ around the t_k epoch, then the illustrated DLD functions result. Assuming a balanced PRN code, each of these events occurs with equal probability. By symmetry we can include the corresponding events where the PRN code value following the t_{k-1} epoch is -1 ; thus, the probability of each of these events is $1/4$. Summing the individual DLD functions results in the EML DLD function for the correlation period as a whole that was

shown in Section 2.3. The signal gate approach also lends itself to computation of DLD functions with filtered and other complex signals, and for determination of noise effects.

Examining the derivation of the EML DLD function illustrated in Fig. 11, it can be seen that the non-zero multipath response between d_1 and $1-d_1$ is due to the rectangular EML gate signal integrating the PRN code signal removed from the chip transition. This response is outside the linear range of the detector and is of marginal value for code tracking.

To eliminate the effect of signals delayed beyond the linear range of the detector, the gating function must have net zero area, as illustrated in Fig. 12. In Fig. 12, ideal gating functions for carrier (prompt correlator) and code tracking are shown—these gates are symmetric and assume infinite bandwidth. Consideration of asymmetric and nonrectangular signal gates can be found in [31][32][33].

There are two general ways that these multipath-mitigating DLD gating functions can be applied, as illustrated in Fig. 10: “Transition-Gated” and “Epoch-Gated”. Similar to the EML gating case, the transition-gated signals occur whenever the PRN code exhibits a transition, whereas for the epoch-gated case, there is a gate at every chip epoch, even when there is no code transition. In the literature, the transition-gated case is known variously as “Double Delta”, “Strobe Correlator” [27], “Multipath Mitigation Technology-A” [28] “High Resolution Correlator” [29][30], and “Pulse Aperture Correlator” [35]. The derivation of the DLD function for the transition and gated cases is illustrated in Fig. 13. It can be seen that the transition-gated DLD function has a correlation “echo” around one chip delay, whereas the epoch-gated case does not, as shown in Fig. 14(a). These DLD functions result in multipath error bound curves as shown in Fig. 14(b). The gated and EML cases all have the same response to short-delay multipath, but the gated correlators eliminate medium-delay multipath. The epoch-gated processing also eliminates multipath around one-chip delay.

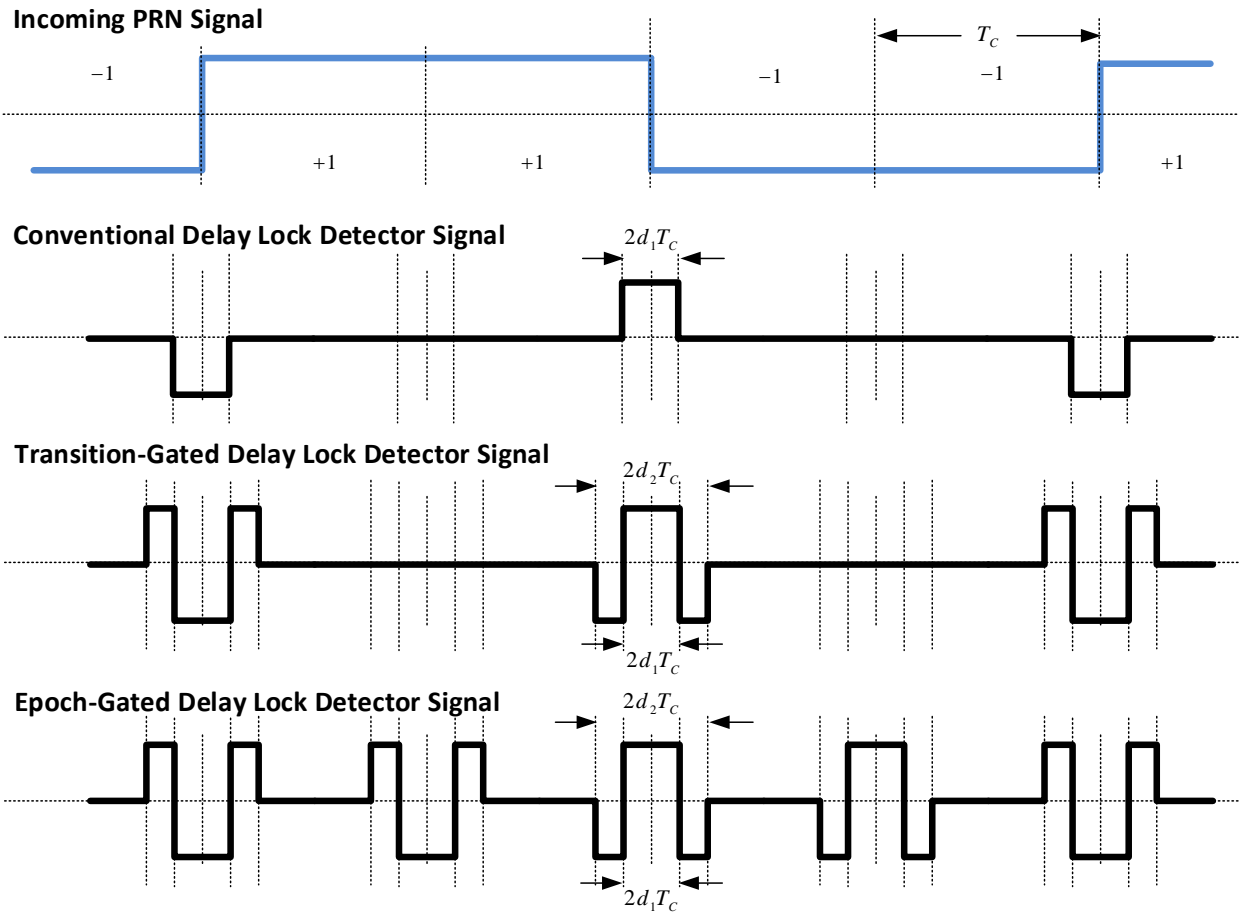


Fig. 10. Illustration of delay lock detector gating functions, with pre-correlation band-limiting neglected.

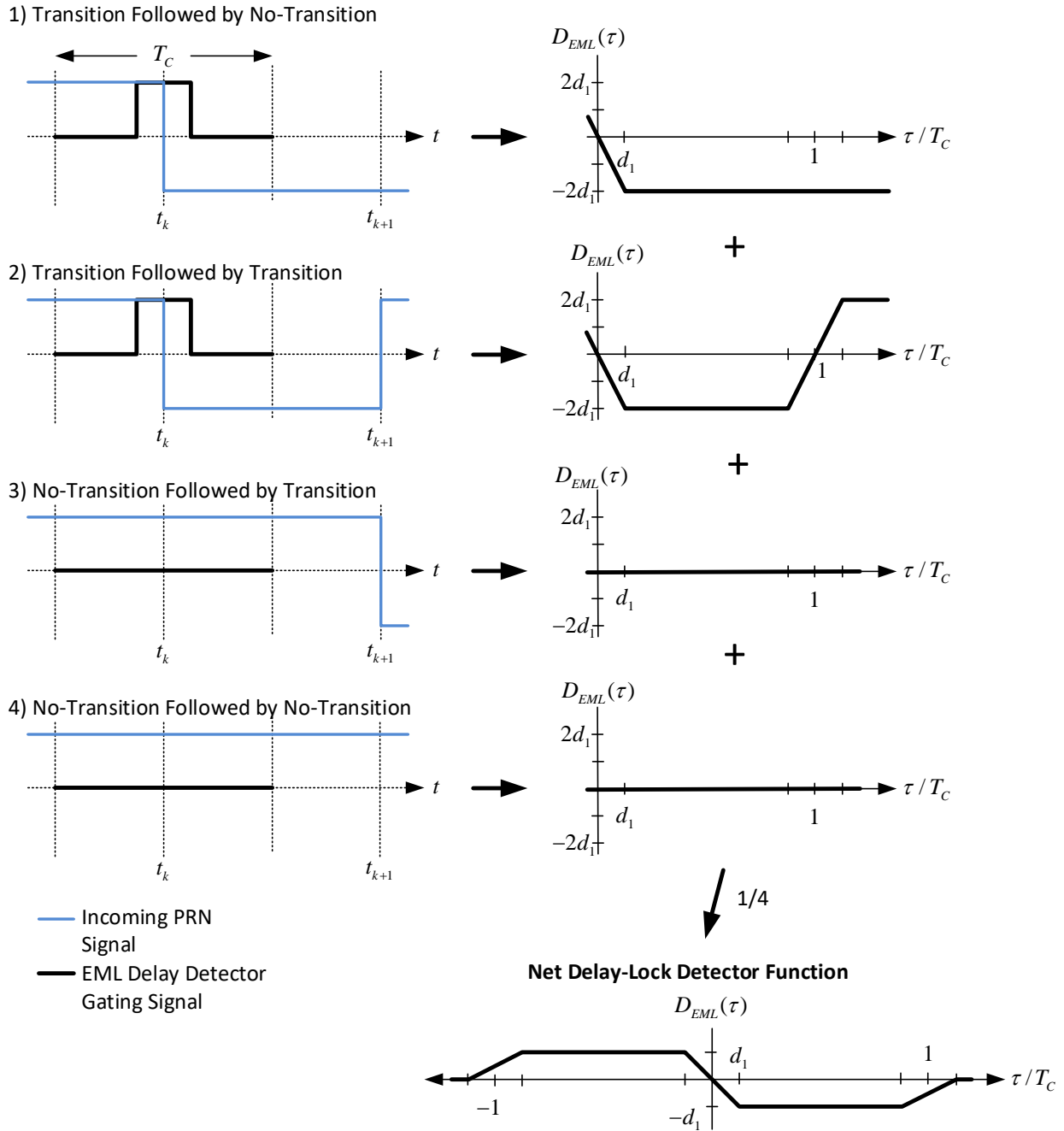


Fig. 11. Derivation of the DLD function for the Conventional (EML) Gate. The left graphs show the incoming and reference signals; the right graphs show corresponding correlation functions, $D_{EML}(\tau)$, computed over a chip period, as a function of the code tracking error, τ .

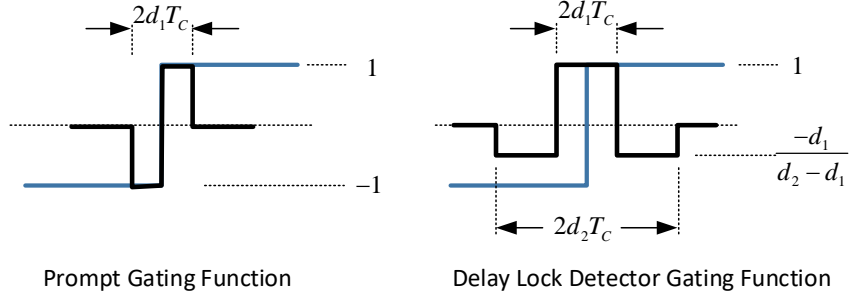


Fig. 12. Ideal symmetric gating functions for enhanced multipath mitigation.

The superior multipath-mitigation performance of transition- and epoch-gating comes at the cost of slightly increased noise, compared to a basic EML gate, as will be discussed in the next section. It should also be noted that the use of epoch- or transition-gating has mostly been used for BPSK(1) signals for high accuracy applications, such as survey. As shown in Section 2.3, BPSK(10) and BOC(1,1) signals have superior multipath characteristics using conventional signal tracking, so the incentive to employ signal gating techniques is not as great.

The use of a prompt signal gate as illustrated in Fig. 12 for mitigation of multipath errors in carrier phase tracking has not been widely used because there is a severe SNR penalty: only the fraction of the incoming signal not zeroed out by the prompt gate is included in the prompt signal coherent integration. For a gate that is one-tenth of a code chip, a 10 dB penalty in SNR is incurred [30]. For code tracking, information about the tracking error is only observable at chip transitions, so there is no loss in signal by narrowing the DLD gate function, as long as the gate width is longer than the duration of the chip transition.

3.2.2 Code Measurement Noise Characteristics

While a reduced width with an EML DLD signal gate offers both improved noise and multipath mitigation, there are noise performance tradeoffs with transition- or epoch-gating. This will be demonstrated by considering the noise portion of (6) in the coherent integration processing illustrated in Fig. 3 with the signal DLD gating function replacing those in Eqn. (7). Ignoring Doppler frequency tracking errors and pre-correlation band-limiting, the DLD noise is

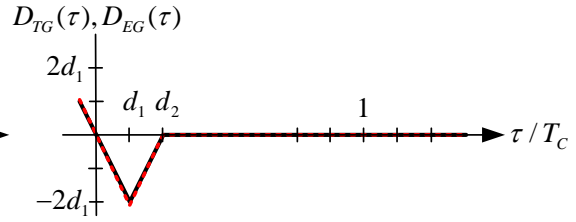
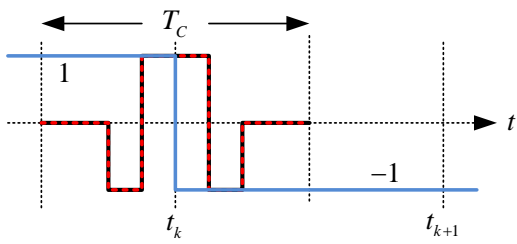
$$w_D = \frac{1}{\sqrt{n_0 T_{PDI}}} \int_0^{T_{PDI}} w(t) s_D(t - \tau) dt \quad (22)$$

Because $w(t)$ is mean-zero WGN, $E\{w_D\} = 0$, the variance can be computed as:

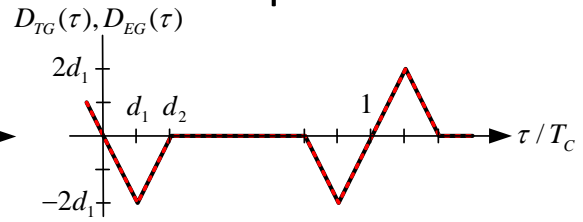
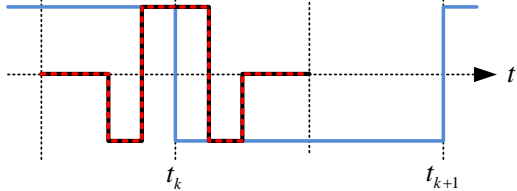
$$\begin{aligned} E\{w_D^2\} &= \frac{1}{n_0 T_{PDI}} \int_0^{T_{PDI}} \int_0^{T_{PDI}} E\{w(t)w(u)\} s_D(t - \tau) s_D(u - \tau) dudt \\ &= \frac{1}{n_0 T_{PDI}} \int_0^{T_{PDI}} \int_0^{T_{PDI}} n_0 \delta(t - u) s_D(t - \tau) s_D(u - \tau) dudt \\ &= \frac{1}{T_{PDI}} \int_0^{T_{PDI}} s_D(t - \tau)^2 dt \end{aligned} \quad (23)$$

where the sifting property of Dirac Delta Function is used.

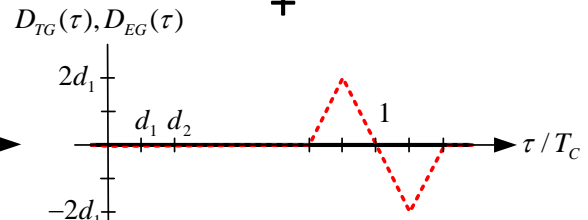
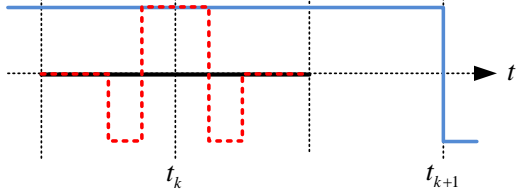
1) Transition Followed by No-Transition



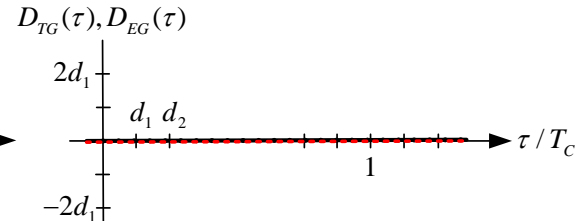
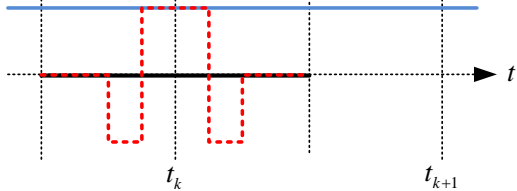
2) Transition Followed by Transition



3) No-Transition Followed by Transition



4) No-Transition Followed by No-Transition



- Incoming PRN Signal
- Chip Transition Delay Lock Gating Signal
- - - Chip Epoch Delay Lock Gating Signal

Net Delay-Lock Detector Functions

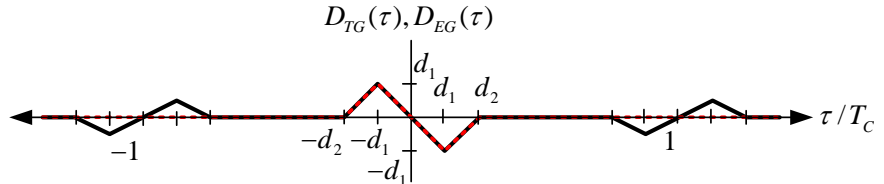


Fig. 13. Derivation of the DLD function for Transition- and Epoch-Gated delay-lock functions. The left graphs show the incoming and reference signals; right graphs show their correlation functions, $D_{TG}(\tau)$ and $D_{EG}(\tau)$, computed over a chip period, as a function of the code tracking error, τ .

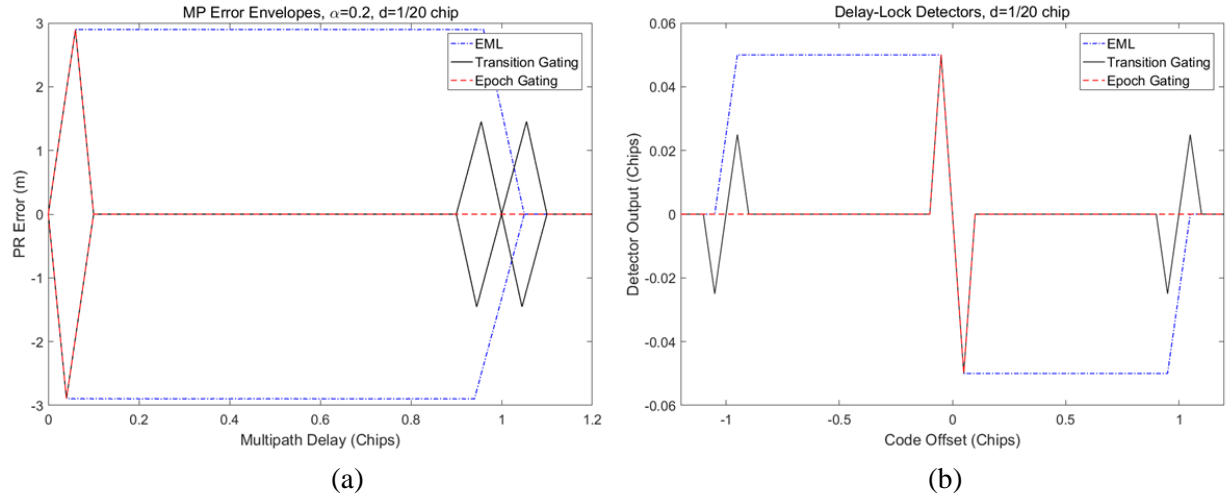


Fig. 14: Comparison of EML Dot Product, Transition Gated and Epoch Gated responses: (a) Delay Lock Detector functions for a BPSK(1) signal with a correlator spacing of 1/20 chip; (b) Corresponding pseudorange multipath error envelopes for a multipath/direct signal ratio $\alpha = 0.2$.

When the gating function is zero, both the signal and noise components of the coherent integration process illustrated in Fig. 11 or Fig. 13 will be blanked. Because the epoch- and transition-gated functions are wider than the equivalent EML gate, the noise will be larger. Equation (23) can be evaluated for the different gate types by considering the non-zero portions of the integration. Over the coherent integration interval there are $N_E = T_{PDI} / T_C$ code epochs and, assuming a balanced PRN code, $N_T = T_{PDI} / 2T_C$ chip transitions (and recalling that the noise variance has been normalized to 1):

EML Gate

$$E\{w_D^2\} = \frac{1}{T_{PDI}} N_T \int_{[2d_1 T_C]} (1)^2 dt = \frac{1}{T_{PDI}} \left(\frac{T_{PDI}}{2T_C} \right) (2d_1 T_C) = d_1 \quad (24)$$

Transition DLD Gate

$$\begin{aligned} E\{w_D^2\} &= \frac{1}{T_{PDI}} N_T \left[\int_{[2d_1 T_C]} (1)^2 dt + 2 \int_{[(d_2 - d_1) T_C]} \left(\frac{d_1}{d_2 - d_1} \right)^2 dt \right] = \frac{1}{T_{PDI}} \left(\frac{T_{PDI}}{2T_C} \right) \left(2d_1 T_C + \frac{2d_1^2 T_C}{d_2 - d_1} \right) \\ &= d_1 + \frac{d_1^2}{d_2 - d_1} \end{aligned} \quad (25)$$

Epoch DLD Gate

$$\begin{aligned} E\{w_D^2\} &= \frac{1}{T_{PDI}} N_E \left[\int_{[2d_1 T_C]} (1)^2 dt + 2 \int_{[(d_2 - d_1) T_C]} \left(\frac{d_1}{d_2 - d_1} \right)^2 dt \right] = \frac{1}{T_{PDI}} \left(\frac{T_{PDI}}{T_C} \right) \left(2d_1 T_C + \frac{2d_1^2 T_C}{d_2 - d_1} \right) \\ &= 2d_1 + \frac{2d_1^2}{d_2 - d_1} \end{aligned} \quad (26)$$

For a typical case of $d_2 = 2d_1$ the transition gate suffers a 3 dB loss compared to the EML, whereas the epoch gate has a 6 dB loss.

3.2.3 Generation of Transition Gates with Conventional Correlators

The transition-gated DLD can be generated with conventional correlator taps [29],[30], unlike epoch-gating which requires explicit generation of the gating function. As a specific example, consider four correlators with adjacent separation of d chips, which are denoted as Early 1 and 2, and Late 1 and 2:

$$\begin{aligned} E2(\tau) &= R(\tau + 3dT_c / 2) \\ E1(\tau) &= R(\tau + dT_c / 2) \\ L1(\tau) &= R(\tau - dT_c / 2) \\ L2(\tau) &= R(\tau - 3dT_c / 2) \end{aligned} \quad (27)$$

The ‘‘High Resolution Correlator 4’’ (HRC4) prompt and DLD functions are synthesized as follows [30]:

$$P_{HRC} = [(E1 + L1) - (E2 + L2)] / 2 \quad (28)$$

$$D_{HRC} = [3(E1 - L1) - (E2 - L2)] / 4 \quad (29)$$

In the context of Fig. 12, (29) corresponds to $d_1 = d/2$ and $d_2 = 3d/2$. For code tracking, the dot-product discriminator uses the HRC4 DLD and a synthesized prompt function:

$$\varepsilon_{DLL} = \frac{ID_{HRC} \cdot IP_0 + QD_{HRC} \cdot QP_0}{IP_0^2 + QP_0^2}, \quad P_0(\tau) = [E1 + L1] / 2 \quad (30)$$

Computer simulation and laboratory test results comparing the HRC4, Eqn. (30), with $d = 0.1$ and EML with $d = 0.05$ chip, for C/A code are shown in Fig. 15. The EML case assumes the presence of a prompt correlator, so the width of the inner portion of the HRC4 gate is the same as the EML gate. The effects of band-limiting are clearly evident.

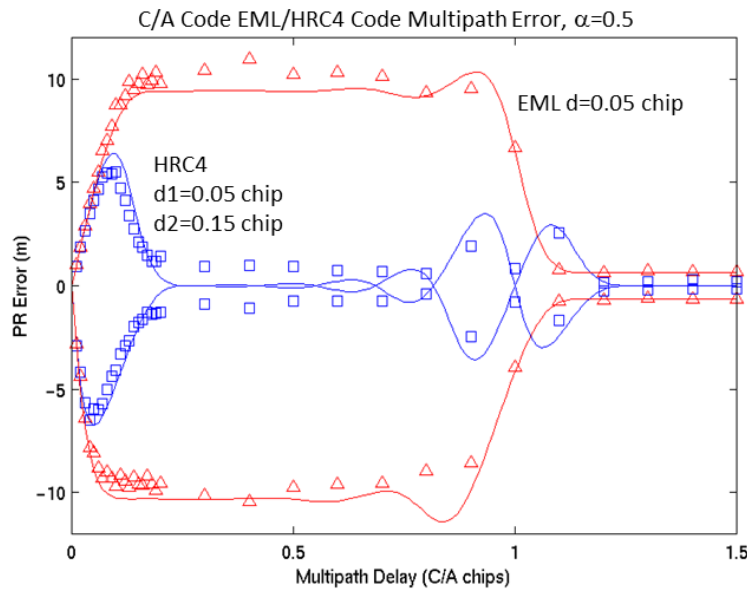


Fig. 15: Simulated (solid lines) and measured (symbols) multipath error envelopes for EML and HRC4 cases with $\alpha_1 = 0.5$ [30]. The measured receiver data was obtained using a GPS simulator with a single multipath signal modeled. The simulation assumed an RF bandpass bandwidth of 16 MHz.

3.3 Parameter Estimation Techniques

The correlator signal processing approaches discussed in the previous section make no claims on optimality. The first attempt to obtain optimal multipath mitigation was the Multipath Estimating Delay Locked Loop (MEDLL) of van Nee, et al. [36]. MEDLL uses a correlator signal model similar to Eqn. (8) as the basis of a Maximum Likelihood Estimation (MLE) approach to estimate signal model parameters $\{\alpha_i, \delta_i, \phi_i\}$, $i = 0, \dots, n$ to provide both code and carrier multipath mitigation. MEDLL uses a bank of narrowly-spaced conventional correlators that sample the autocorrelation function for both in-phase and quadra-phase. This can require tens of correlators per tracking channel—a significant hardware cost. MEDLL has been demonstrated to be effective when there are only one or two dominant multipath signals with sufficiently long delays to enable observability.

Building on MEDLL and correlation signal processing techniques, Weill proposed the use of “super-resolution techniques” [37],[38],[39], wherein the local code reference signals are matched to the time-domain characteristics of the pre-correlation signals. This permits discrimination of short delay multipath, subject only to information-theoretic bounds due to signal bandwidth. This approach has been coupled with MLE signal parameter estimation in the so-called Vision Correlator [40] to achieve near theoretical error bounds when the signal conditions match model assumptions. A review of multi-correlator-based parameter estimation techniques may be found in [41].

There are a number of significant drawbacks to parameter estimation techniques. First, they are computationally burdensome, particularly as the number of multipath replicas modeled increases. Super-resolution processing requires specialized hardware, which much be matched to the receiver characteristics. Lastly, there is the problem of model fidelity. If the number of multipath signals is under- or over-estimated, then MLE techniques can perform poorly. MLE also can have convergence problems in poor SNR. For these reasons, parameter estimation techniques have typically been applied to fixed-site applications that may have a single-dominant reflection source to be eliminated.

3.4 Array-Based Signal Processing

The ultimate method of multipath mitigation is to use multiple antenna elements and adaptive array processing. Ray, et al. [42] demonstrated code and carrier phase multipath mitigation by using an array of half-wavelength spaced antenna elements. They assumed that, over a given observation interval, the multipath returns could be represented by a single multipath reflector. The signals from the multiple antennas were used to separate out the direct signal from the reflected signal and the code and carrier measurements extracted from the direct signal estimate.

Many adaptive array anti-jam systems use pre-correlation digital beam forming (DBF) to synthesize a beam in the direction of arrival of the satellite signal (using known platform attitude from an inertial measurement unit) while also nulling jammers. Such systems have demonstrated multipath mitigation capabilities [43],[44]. The digital beam steering theoretically enhances the direct signal power proportional to the number of antenna elements. So for seven elements, this is about an 8 dB reduction in α_i , which in many circumstances could be significant.

Pre-correlation DBF does nothing to explicitly mitigate multipath signals (i.e., the multipath mitigation benefits are incidental). DBF processing can also be done post-correlation [44]. In this case the multipath signals integrate above thermal noise, permitting the DBF algorithm to detect and actually generate a null

in the spatial direction of the multipath. This type of processing has not been widely deployed due to the size and cost of antenna arrays, separate RF front end and analog-to-digital converters (ADC) for each antenna, and tracking channels for each antenna/satellite combination. Also an increased number of ADC bits may be required in order to maintain dynamic range. However, the increasing density and clock speeds of integrated circuits is driving down the cost of the required signal processing components needed for post-correlation DBF, so it is expected that this technology will become commercially available in the 2020s.

3.5 Doppler Domain and Synthetic Aperture Techniques

When the receiver is moving, different components of a multipath-contaminated signal can have different Doppler shifts. Integrating the correlator outputs coherently over 100 ms or more provides sufficient resolution at road vehicle speeds to separate the different signal components in the Doppler domain, preventing reflected signals from interfering with the direct LOS signal [45][46]. Note that this does not work for signals reflected off surfaces parallel to the direction of travel as these have the same Doppler shift as the direct LOS signal.

This technique is equivalent to beamforming with a synthetic aperture, a common radar processing technique, maximizing the gain in the direction of the direct LOS signal and, in general, reducing it in the direction of the reflected signals [47][48]. If the receiver is travelling in a straight line, the gain pattern will be symmetric about the direction of travel. Changes of direction during the coherent integration period enable an asymmetric gain pattern. The longer the coherent integration period, the sharper the beam formed, but the more accurate the frequency control of the carrier NCO must be to maintain coherence. This typically requires either correlating at multiple Doppler shifts in parallel or providing Doppler aiding, either from other GNSS signals or from external aiding sources, such as inertial sensors.

4. Carrier Smoothed Code

Code pseudorange noise and multipath errors are on the order of 100 times larger than corresponding carrier phase errors. This is the motivation for use of carrier phase in precision positioning applications discussed in Chapters 19 and 20. However, carrier phase positioning typically requires external augmentation or reference receiver data sources, so most stand-alone GNSS positioning ultimately relies on pseudorange measurements. The smoothing of GNSS code pseudorange measurements with carrier phase measurements was introduced by Hatch [49] and is now a well-established GNSS signal processing technique. Note that carrier smoothing does nothing to mitigate errors due to NLOS reception.

4.1 Carrier Smoothing Formulations

For the purposes of the carrier smoothing filter derivation and performance modeling, we use the following simplified models for the GNSS code and carrier phase measurements at frequency f_L from a single transmitter:

$$\begin{aligned}\rho_L &= r + \delta_T + \delta_R + I_L + T + \delta\rho_{ML} + \varepsilon_{\rho L} \\ \varphi_L &= r + \delta_T + \delta_R - I_L + T + \delta\varphi_{ML} + \varepsilon_{\varphi L} + N_L\lambda_L\end{aligned}\tag{31}$$

where

ρ_L = Code pseudorange measurement (in meters)

φ_L = Carrier phase measurement (in meters)

r = Geometric Line-of-Sight (LOS) range

δ_T = Satellite clock and ephemeris errors projected along LOS

δ_R = Receiver clock bias

I_L = Ionospheric refraction at f_L

T = Tropospheric refraction

$\delta\rho_{ML}, \delta\varphi_{ML}$ = Code and carrier multipath at f_L

$\varepsilon_{\rho L}, \varepsilon_{\varphi L}$ = Code and carrier receiver noise and other errors

$N_L\lambda_L$ = Carrier phase ambiguity for the carrier with wavelength λ_L , where N_L is an integer.

When the development does not require denoting multiple carrier frequencies, then the subscript L may be omitted.

The carrier phase ambiguity complicates using the carrier for positioning. However, carrier phase differences can provide a very accurate measurement of the change in pseudorange for non-dispersive effects (LOS range, clocks and troposphere). The idea behind carrier smoothing is to combine carrier phase and code measurements to obtain a pseudorange that has the low frequency characteristics of the code measurements, but with the high frequency errors of the phase measurements.

There are two common approaches for carrier smoothing: the Hatch Filter [26] and the Complementary Filter [50], illustrated in Fig. 16(a) and (b), respectively. The block diagrams in Fig. 16 are discrete-time signal processing/estimation representations, where z^{-1} is the unit delay operator, i.e., $z^{-1}x(t_n) = x(t_{n-1})$, and $x(t_n^-)$ represents an extrapolated value valid at time step t_n based on the past data $\{t_1, \dots, t_{n-1}\}$. It is a straightforward algebraic exercise to show that these formulations are equivalent. The Hatch filter directly produces a smoothed pseudorange, whereas the complementary filter first smooths the code-minus-carrier (CMC) signal, $\chi(t_n)$, then adds the carrier phase back on to obtain the smoothed pseudorange. A common choice for the gain at time $t_n, n = 0, 1, \dots$, is:

$$K_n = \begin{cases} 1/n, & n = 1, \dots, N_{\max} - 1 \\ 1/N_{\max}, & n \geq N_{\max} \end{cases} \quad (32)$$

This time-varying gain sequence will be shown to possess certain optimal properties. The determination of N_{\max} in terms of filter bandwidth or smoothing time constant will be discussed below. With this choice, the Hatch filter from Fig. 16(a) can be written compactly as:

$$\bar{\rho}(t_n) = \begin{cases} \frac{n-1}{n} [\bar{\rho}(t_{n-1}) + \Delta\varphi(t_n)] + \frac{1}{n} \rho(t_n), & n = 1, \dots, N_{\max} - 1 \\ \frac{N_{\max} - 1}{N_{\max}} [\bar{\rho}(t_{n-1}) + \Delta\varphi(t_n)] + \frac{1}{N_{\max}} \rho(t_n), & n \geq N_{\max} \end{cases} \quad (33)$$

There are pros and cons for either implementation. The Hatch filter uses carrier phase delta-range instead of carrier phase, avoiding estimation of a CMC bias, which may facilitate filter restart after cycle slips. On the other hand, the complementary filter can be readily coasted during measurement outages, since $\chi(t_n)$ evolves slowly, even on moving platforms.

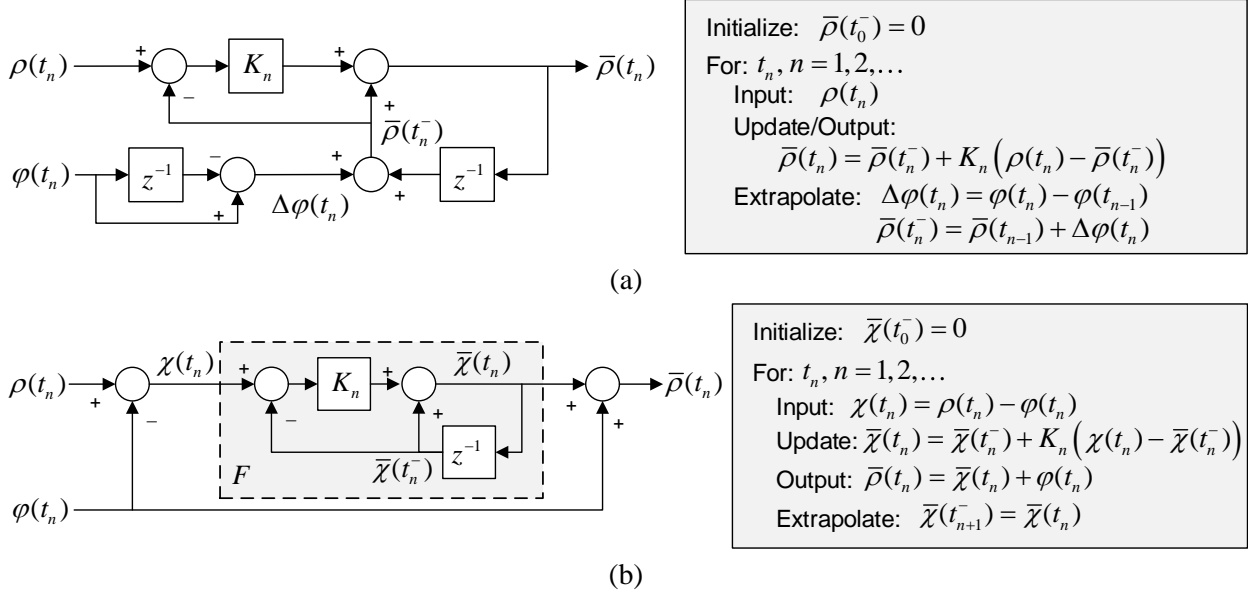


Fig. 16. Carrier smoothing block diagrams and filter processing equations: (a) Hatch filter; (b) Complementary filter.

The qualitative behavior of carrier smoothing will be demonstrated using the complementary filter. The filtering operation in the shaded box in Fig. 16(b) can be denoted as an operator, F . The overall input-output relation is:

$$\bar{\rho} = F(\rho - \varphi) + \varphi = F\rho + (1 - F)\varphi \quad (34)$$

For the case of a fixed- or steady-state gain, $K_n \equiv K$, and the complementary filter iterative equations for F in Fig. 16(b) can be written as

$$\bar{\chi}(t_n) = (1 - K)\bar{\chi}(t_{n-1}) + K\chi(t_n) \quad (35)$$

This discrete time equation can be written in terms of a Z-transform [53] as

$$F(z) = \frac{K}{1 - (1 - K)z^{-1}} = \frac{Kz}{z - (1 - K)} \quad (36)$$

This will be a low-pass filter for $0 < K < 1$, which is satisfied by the gain sequence (32), and implies $1 - F$ will be high-pass. Using the measurement model (31) in (34) yields:

$$\bar{\rho}_L = (r + \delta_T + \delta_R + T) + (2F - 1)I_L + F(\delta\rho_{ML} + \varepsilon_{\rho L}) + (1 - F)(\delta\varphi_{ML} + \varepsilon_{\varphi L} + N_L\lambda_L) \quad (37)$$

Note that the LOS range, clock and tropospheric terms are unfiltered, so the desired pseudorange measurement components are unaffected by smoothing. The pseudorange noise and multipath will be low-pass filtered, whereas the carrier ambiguity, noise and multipath are high-pass filtered. Since at zero frequency ($z = 1$), $F(1) = 1$, the smoothed pseudorange will have the same long-term error trends due to low-frequency multipath and other errors as the unsmoothed measurement. The filtered ionospheric term will cause a bias in the measurement that is proportional to the rate of change of the ionospheric delay [52].

4.2 Smoothing Filter Gain

4.2.1 Kalman Filter Derivation

The smoothing filter gain, K_n can be determined from a Kalman Filter formulation [50],[51]. From (31), the CMC measurement, $\chi(t_n)$, is

$$\chi = 2I + (M_\rho + \varepsilon_\rho) - (M_\varphi + \varepsilon_\varphi + N \lambda) \quad (38)$$

The carrier ambiguity will be a constant bias as long as there are no cycle slips. As a limiting case, consider a simplified model of (38) consisting of a slowly varying bias plus WGN:

$$\chi(t_n) = \chi_0(t_n) + \varepsilon(t_n) \quad (39)$$

The WGN $\varepsilon(t_n)$ measurement noise is mostly due to code pseudorange noise, whose covariance we will denote as R_ε . The bias term will be modelled as a Gauss-Markov process:

$$\chi_0(t_{n+1}) = \beta \chi_0(t_n) + w_\chi(t_n) \quad (40)$$

Here w_χ is WGN, with $\text{cov } w_\chi = Q$. With the process model (40) and measurement model (39) we can develop a Kalman Filter for the CMC bias. Denote $\bar{\chi}(t_n) = E\{\chi_0(t_n) | \chi(t_k), k=1, \dots, n\}$ with covariance $P(t_n)$ and $\bar{\chi}(t_n^-) = E\{\chi_0(t_n) | \chi(t_k), k=1, \dots, n-1\}$ with covariance $P(t_n^-)$. Typically, $P(t_1^-)$ is chosen so $P(t_1^-) \gg R_\varepsilon$ to represent ignorance of the CMC value prior to the first measurements being processed. The Kalman Filter equations for $n=1, 2, \dots$ are:

Measurement Update

$$K_n = \frac{P(t_n^-)}{P(t_n^-) + R_\varepsilon} \quad (41)$$

$$\bar{\chi}(t_n) = (1 - K_n) \bar{\chi}(t_n^-) + K_n \chi(t_n), \quad P(t_n) = (1 - K_n) P(t_n^-) = \frac{R_\varepsilon P(t_n^-)}{P(t_n^-) + R_\varepsilon}$$

Extrapolate

$$\bar{\chi}(t_{n+1}^-) = \beta \bar{\chi}(t_n), \quad P(t_{n+1}^-) = \beta^2 P(t_n) + Q \quad (42)$$

The steady-state bandwidth of the smoothing filter can be set by varying the Gauss-Markov model parameters. Note that as $P(t_1^-) \rightarrow \infty$, $K_1 \rightarrow 1$ and $P(t_1) \rightarrow R_\varepsilon$, which says the initial smoothed pseudorange error will be bounded by the unsmoothed noise.

The filter gain sequence in Eq. (32) is a special case of (41), with $\beta=1$ and $Q=0$, which will now be demonstrated. With these choices for β and Q , and combining the update and extrapolation steps, an iterative expression for the smoothed covariance is:

$$P(t_n) = P(t_{n-1}) - \frac{P(t_{n-1})^2}{P(t_{n-1}) + R_\varepsilon} = \frac{P(t_{n-1})R_\varepsilon}{P(t_{n-1}) + R_\varepsilon} \quad (43)$$

A solution to (43) is

$$P(t_n) = \frac{R_\varepsilon}{n} \quad (44)$$

This is proven by noting that

$$P(t_{n+1}) = \frac{P(t_n)R_\varepsilon}{P(t_n) + R_\varepsilon} = \frac{R_\varepsilon^2 / n}{R_\varepsilon / n + R_\varepsilon} = \frac{R_\varepsilon}{n+1} \quad (45)$$

which is consistent with (44), thus demonstrating this assertion by induction.

With the expression (44) for the smoothed covariance, the filter gain can be written using (41) as

$$K_n = \frac{P(t_{n-1})}{P(t_{n-1}) + R_\varepsilon} = \frac{R_\varepsilon / (n-1)}{R_\varepsilon / (n-1) + R_\varepsilon} = \frac{1}{n} \quad (46)$$

This shows that the gain sequence in Eq. (32) is optimal for smoothing a constant bias in WGN. In practice, the assumption that the CMC bias is constant may only be valid over limited intervals, therefore the floor on the gain in (32) for time steps beyond N_{\max} is used to set the steady-state smoothing time constant.

4.2.2 Steady-State Gain Derivation

The value for N_{\max} can be determined by relating the CMC filter F to a first-order, continuous-time, low-pass filter:

$$F(s) = \frac{1}{T_0 s + 1} \quad (47)$$

In (47), T_0 is the time constant of the filter. For a sampling interval ΔT , a matched pole-zero Z-transform of $F(s)$ is given by [53]

$$F(z) = \frac{(1 - e^{-\Delta T/T_0})z}{z - e^{-\Delta T/T_0}} \quad (48)$$

Equating the terms in (48) with the transfer function in (36) and the gain expression in (46) yields

$$\begin{aligned} K &= \frac{1}{N_{\max}} = 1 - e^{-\Delta T/T_0} \approx \Delta T / T_0, \quad \Delta T \ll T_0 \\ \Rightarrow N_{\max} &\approx T_0 / \Delta T \end{aligned} \quad (49)$$

where T_0 is the time constant of the filter and ΔT is the sampling interval.

Selecting N_{\max} using (49) yields a fixed gain filter for $n > N_{\max}$ whose covariance equation can be written as [51]

$$P_{n+1} = (1 - K)^2 P_n + K^2 R_\varepsilon \quad (50)$$

In steady-state, $P_n \rightarrow P_s$, $n \rightarrow \infty$. Then (50) becomes

$$\begin{aligned} [1 - (1 - K)^2] P_s &= K^2 R_\varepsilon \\ \Rightarrow P_s &= \frac{K R_\varepsilon}{2 - K} \approx \frac{R_\varepsilon}{2 N_{\max}} \approx \frac{\Delta T}{2 T_0} R_\varepsilon, \quad T_0 \gg \Delta T \end{aligned} \quad (51)$$

Note that this is only valid if the pseudorange measurement noise is statistically independent from sample to sample. This can be guaranteed by selecting the measurement output rate to be slower than the noise equivalent bandwidth of the code tracking loop. Also, (51) ignores the contribution of the carrier phase

errors on the output of the low-pass filter in Fig. 16(b), so a more accurate expression for the steady-state smoothed pseudorange error is

$$P_s = \frac{KR_\epsilon}{2-K} + \sigma_\phi^2 \quad (52)$$

where σ_ϕ^2 is the carrier phase noise variance.

For specular multipath, with a fading frequency δf_i , then the attenuation of the pseudorange multipath error can be approximated using (47):

$$|F(j2\pi\delta f_i)| = \frac{1}{|1+sT_0|} \Big|_{s=j2\pi\delta f_i} = \frac{1}{|1+j2\pi\delta f_i T_0|} = \frac{1}{\sqrt{1+(2\pi\delta f_i T_0)^2}} \quad (53)$$

To achieve significant multipath attenuation, we must have $T_0 \gg 1/2\pi\delta f_i$. For example, consider the slow fading experienced at a fixed site, as described in Section 2.1, with $\delta f_i \approx 1.5$ mHz. To achieve a factor of five reduction in multipath amplitude requires $T_0 > 520$ s, illustrating the challenge of mitigating strong multipath at fixed sites.

4.3 Ionospheric Errors and Dual-Frequency Smoothing

4.3.1 Ionospheric Divergence Errors in Single Frequency Carrier Smoothing

The ultimate usefulness of carrier smoothing for GNSS navigation processing, particularly in single frequency local differential GNSS processing, is dependent on the characteristics of the filtered ionospheric term in Eqn. (37). Over the time-scale of the carrier smoothing, the change in the ionospheric refraction can be approximated as bias plus a ramp:

$$I_L(t) = I_0 + I_d t \quad (54)$$

This can be written in the Laplace Transform “s” domain as

$$I_L(s) = \frac{I_0}{s} + \frac{I_d}{s^2}$$

As can be seen in Equations (31) and (37), the change in the ionospheric error causes a divergence between GNSS code and carrier measurements. The error in the ionospheric residual induced by the smoothing filter can be determined using (47) by

$$\begin{aligned} \Delta I_L &= I_L - (2F(s) - 1)I_L = 2[1 - F(s)]I_L \\ &= \frac{2T_0 s}{T_0 s + 1} I_L \end{aligned} \quad (55)$$

The steady-state behavior of ΔI_L may be determined by applying the final value theorem of Laplace Transforms

$$\Delta I_{ss} = \lim_{s \rightarrow 0} s \cdot \left[\frac{2T_0 s}{T_0 s + 1} \left(\frac{I_0}{s} + \frac{I_d}{s^2} \right) \right] = 2T_0 I_d \quad (56)$$

Equation (56) confirms that the carrier smoothing filter introduces an error as it tracks a time-varying ionospheric delay ($I_d \neq 0$), the size of which is proportional to the smoothing time constant T_0 . For a

typical value of ionospheric divergence of 10 cm/min and 100 s smoothing, the offset is about 33 cm, which would be a non-negligible error for local-area DGNS applications, such as aircraft precision approach and landing (Chapter 12). During a severe ionospheric storm, the divergence can be 5-7 times as large, thus it is critical in single-frequency DGNS applications for the smoothing filters in the reference and rover receivers to use the same time constant, so that the ionospheric divergence error cancels when the measurements between the rover and reference are differenced.

4.3.2 Divergence-Free Filter Derivation

The ionospheric divergence effect in carrier smoothing can be addressed by using dual-frequency measurements. Following the approach in [54], we simplify the measurement model (31) by lumping terms that are affected by carrier smoothing in similar ways:

$$\begin{aligned}\rho_L &= R + I_L + w_{\rho L} \\ \varphi_L &= R - I_L + w_{\varphi L} + N_L \lambda_L\end{aligned}\quad (57)$$

where:

$$\begin{aligned}R &= r + \delta_T + \delta_R + T \\ w_{\rho L} &= \delta\rho_{ML} + \varepsilon_{\rho L} \\ w_{\varphi L} &= \delta\varphi_{ML} + \varepsilon_{\varphi L}\end{aligned}\quad (58)$$

The ionospheric refraction will be modeled as

$$I_L = \frac{K_L}{f_L^2}\quad (59)$$

Define generalized code and carrier phase measurements by combining the code and carrier measurements from frequencies f_1 and f_2 as:

$$\begin{aligned}\rho &= a_1\rho_1 + a_2\rho_2 \\ \varphi &= b_1\varphi_1 + b_2\varphi_2\end{aligned}\quad (60)$$

where a_1 , a_2 , b_1 , and b_2 are coefficients to be determined. To preserve the LOS and clock information in the generalized measurements, we must have:

$$a_1 + a_2 = 1\quad (61)$$

$$b_1 + b_2 = 1\quad (62)$$

Combining (57) and (60), and applying (61) and (62), the CMC can be written as

$$\begin{aligned}\chi &= \rho - \varphi = (a_1\rho_1 + a_2\rho_2) - (b_1\varphi_1 + b_2\varphi_2) \\ &= (a_1 + a_2)R - (b_1 + b_2)R + (a_1w_{\rho 1} + a_2w_{\rho 2}) - (b_1w_{\varphi 1} + b_2w_{\varphi 2}) - (b_1\lambda_1N_1 + b_2\lambda_2N_2) \\ &\quad + (a_1 + b_1)I_1 + (a_2 + b_2)I_2 \\ &= (a_1w_{\rho 1} + a_2w_{\rho 2}) - (b_1w_{\varphi 1} + b_2w_{\varphi 2}) - (b_1\lambda_1N_1 + b_2\lambda_2N_2) + [(a_1 + b_1)I_1 + (a_2 + b_2)I_2]\end{aligned}\quad (63)$$

To eliminate the effect of ionospheric divergence on the smoothing processing, the last term in (63) must be zero. Using (59) we obtain the following condition:

$$(a_1 + b_1)f_2^2 + (a_2 + b_2)f_1^2 = 0 \quad (64)$$

We now have three equations, (61), (62), and (64), for the four unknowns, a_1 , a_2 , b_1 , and b_2 . Therefore one of the parameters can be arbitrarily chosen. We will use a_1 as this independent variable. With this choice we have the following expressions:

$$\begin{aligned} a_2 &= 1 - a_1 \\ b_1 &= \frac{2f_1^2}{f_1^2 - f_2^2} - a_1 \\ b_2 &= a_1 - \frac{f_1^2 + f_2^2}{f_1^2 - f_2^2} \end{aligned} \quad (65)$$

Note that this derivation can be extended to three frequencies and mixed code-carrier combinations [55] which enables additional, simultaneous constraints to be imposed, e.g., maintaining integer ambiguities, while minimizing noise and/or ionospheric errors.

4.3.3 Divergence-Free Measurement Combinations

We now examine some specific special divergence-free cases based on particular choices for α_1 in (65).

Single-Frequency Pseudorange Divergence-Free Smoothing

Choosing $a_1 = 1$ or $a_1 = 0$, yields the divergence-free smoothed f_1 or f_2 pseudorange, respectively, discussed in [52] and [54]. The smoothing filter inputs in these cases are:

$$\begin{aligned} f_1: \quad \rho &= \rho_1, \quad \varphi = \frac{f_1^2 + f_2^2}{f_1^2 - f_2^2} \varphi_1 - \frac{2f_2^2}{f_1^2 - f_2^2} \varphi_2 \\ f_2: \quad \rho &= \rho_2, \quad \varphi = \frac{2f_1^2}{f_1^2 - f_2^2} \varphi_1 - \frac{f_1^2 + f_2^2}{f_1^2 - f_2^2} \varphi_2 \end{aligned} \quad (66)$$

These combinations will have the pseudorange multipath and noise errors associated with the respective carrier frequency, whereas the carrier phase errors are slightly elevated by combining multiple frequency signals.

Ionosphere-Free Smoothing

A special case of divergence-free processing is obtained by eliminating the ionosphere in both ρ and φ individually, rather than just in the CMC as accomplished in (64). From (59) and (60) this is achieved in the generalized code measurement, ρ , by having

$$\frac{a_1}{f_1^2} + \frac{a_2}{f_2^2} = 0 \quad (67)$$

With (61) and (65) we get

Iono-Free:

$$\rho_{IF} = \frac{f_1^2}{f_1^2 - f_2^2} \rho_1 - \frac{f_2^2}{f_1^2 - f_2^2} \rho_2, \quad \varphi_{IF} = \frac{f_1^2}{f_1^2 - f_2^2} \varphi_1 - \frac{f_2^2}{f_1^2 - f_2^2} \varphi_2 \quad (68)$$

Note that this measurement combination suffers from a significant amplification of noise and multipath errors, since the coefficients on the code and carrier measurements are greater than one.

Wide-Lane/Narrow-Lane Combinations

As discussed in [54], the well-known wide-lane and narrow-lane carrier phase combinations have the following divergence-free combinations:

Wide-Lane Carrier Phase/Narrow-Lane Code:

$$\rho_{NL} = \frac{f_1}{f_1 + f_2} \rho_1 + \frac{f_2}{f_1 + f_2} \rho_2, \quad \varphi_{WL} = \frac{f_1}{f_1 - f_2} \varphi_1 - \frac{f_2}{f_1 - f_2} \varphi_2 \quad (69)$$

Narrow-Lane Carrier Phase/Wide-Lane Code:

$$\rho_{WL} = \frac{f_1}{f_1 - f_2} \rho_1 - \frac{f_2}{f_1 - f_2} \rho_2, \quad \varphi_{NL} = \frac{f_1}{f_1 + f_2} \varphi_1 + \frac{f_2}{f_1 + f_2} \varphi_2$$

As shown in [54], the narrow-lane code combination provides attenuation of pseudorange multipath and noise, since the coefficients on the code measurements are less than one.

4.4 Performance Analysis and Examples

In this section we provide examples of carrier smoothing performance with simulated and live data.

Figure 17 shows single frequency carrier smoothing results with GPS L1 P(Y) code in a high multipath rooftop environment. Raw and smoothed CMC are plotted, which illustrates the effect of ionospheric divergence. The unsmoothed data clearly exhibits oscillatory effects of multipath due to satellite motion. A modest smoothing time constant of 100 s provides significant attenuation of the multipath, but residual errors are still evident. A longer time constant of 600 s provides additional attenuation of multipath oscillations, but suffers from the bias associated with the filtered ionospheric divergence as in Eqn. (56).

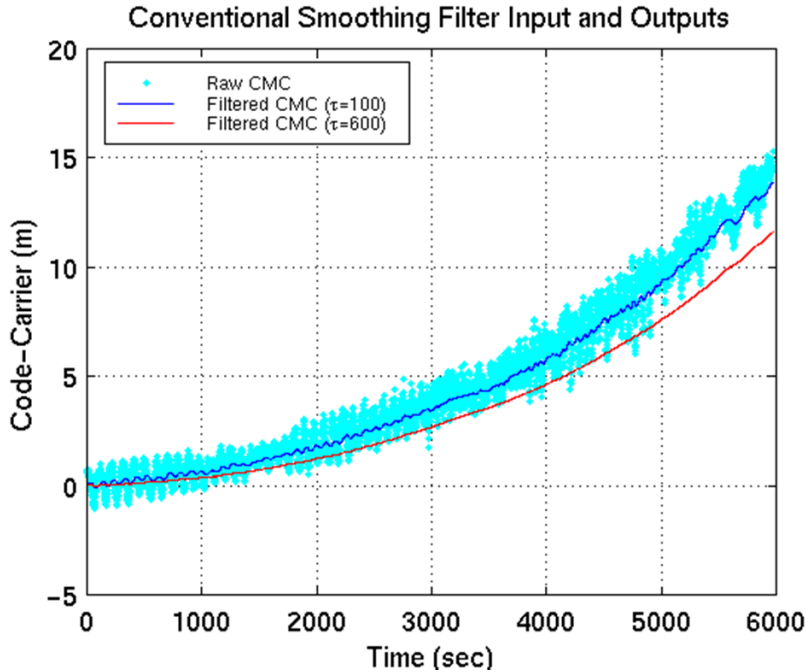


Fig. 17. Single frequency smoothing live data results [52]. The bias induced by ionospheric divergence on the smoothed code measurements can be seen by the offsets in the red and blue curves from raw CMC.

Figure 18 shows the corresponding L1 P(Y)-code divergence-free CMC using the L1/L2 P(Y) carrier phase as in (66). The effects of ionospheric divergence are eliminated, which enables longer smoothing time constants to be used in GNSS/DGNSS applications. Figure 19 shows the corresponding ionospheric-free smoothing results, as in (68), using L1/L2 P(Y) code. The amplification of the multipath and noise compared to the single frequency pseudorange errors in Fig. 18 is evident by the 3X larger peak-to-peak level of the raw CMC.

Figures 20 and 21 show simulated smoothing filter results that illustrate the benefit of the time-varying gain sequence (32) versus using the steady-state gain value in terms of filter convergence. Figure 20 shows the result of a covariance analysis with WGN-only as in (43) with $R_e = (0.5 \text{ m})^2$, carrier phase noise standard deviation of 0.018 cycles, and steady-state time constant $\tau = 100 \text{ s}$. Single-frequency (SF) and L1 divergence-free (DF) smoothing results are depicted. The fixed gain filter has a long convergence time, on the order of two filter time constants, whereas the time-varying Kalman Filter (KF) gain converges quickly for both the single-frequency or divergence-free cases. In steady-state, the divergence-free filter has a slightly elevated error due to the amplification of the carrier phase errors in Eqn. (66), consistent with the analytical result in (52). However, this additional error would likely be swamped by residual multipath and single frequency ionospheric-divergence errors in practice.

Figure 21 illustrates the improved convergence of divergence-free smoothing with a time-varying gain in ionospheric divergence relative to single-frequency smoothing. The simulated initial ionospheric bias at L1 is 4 m with a divergence of 1 cm/s. For single-frequency smoothing, both the fixed- and Kalman-gain have a slow convergence to the biased steady-state trend. In contrast, the divergence-free filter converges very quickly to the unbiased trend.

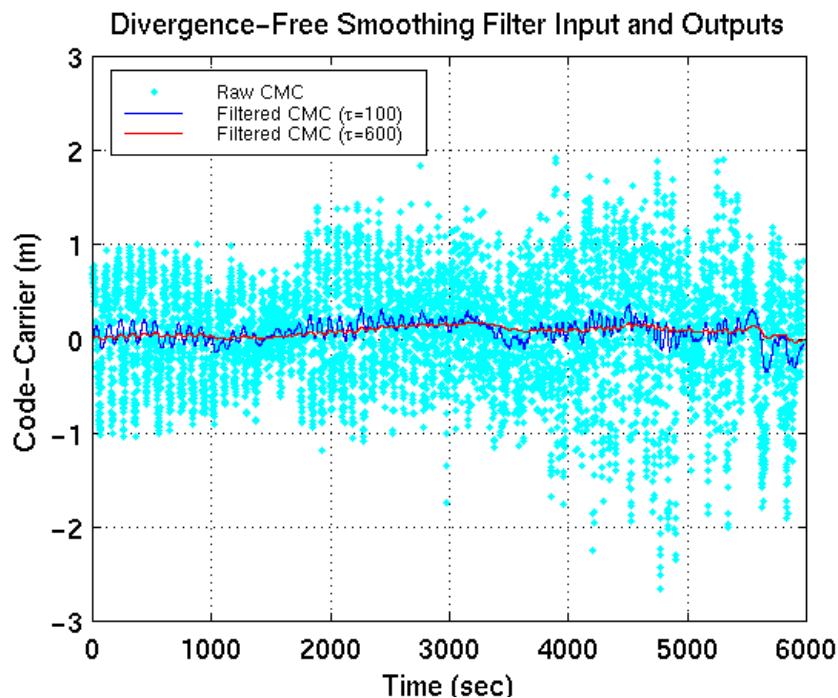


Fig. 18. Divergence-free smoothing live data results [52]. Ionospheric divergence effects on the smoothed code measurement have been eliminated, enabling longer smoothing time constants to be used.

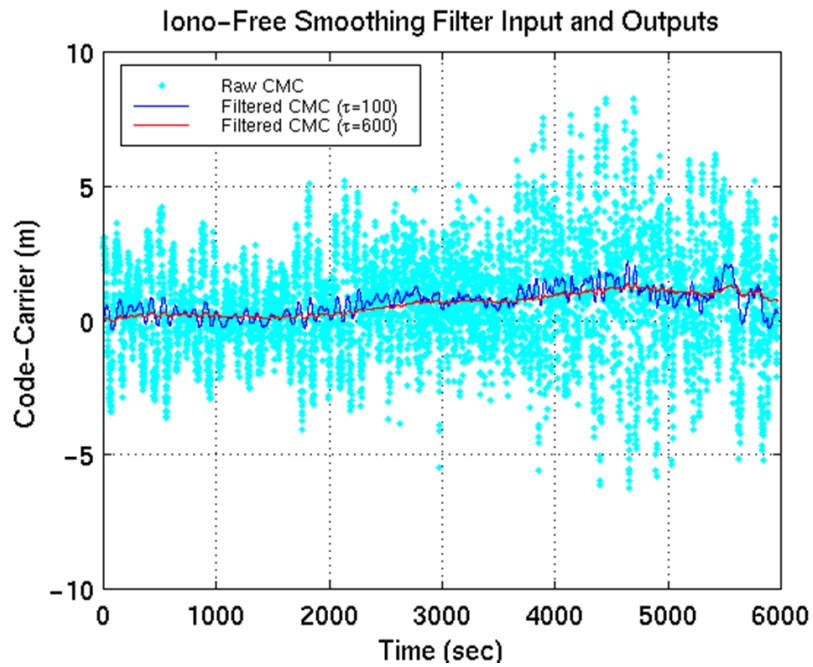


Fig. 19. Ionospheric free smoothed live data results [52]. The effects of ionospheric divergence are eliminated but with a magnification of errors.

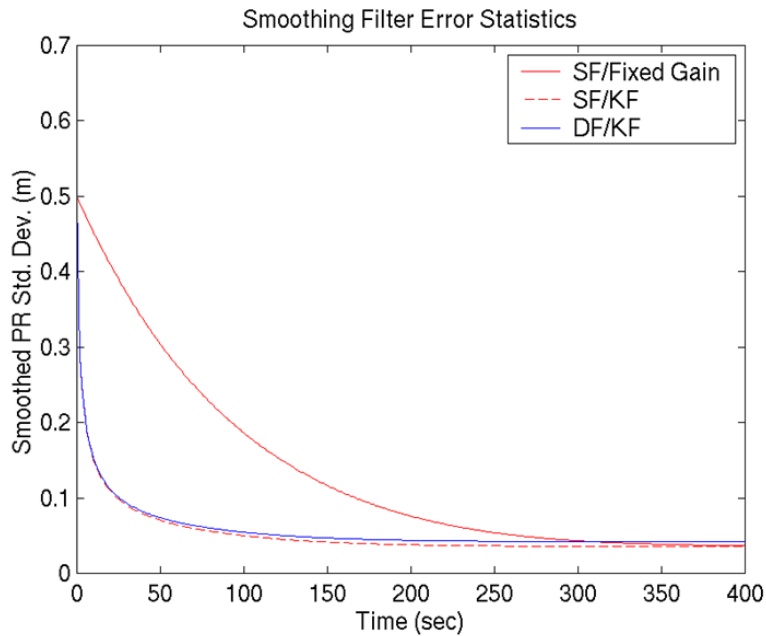


Fig. 20. Covariance analysis simulation of smoothing filter convergence for steady-state time constant of 100 s for smoothing WGN.

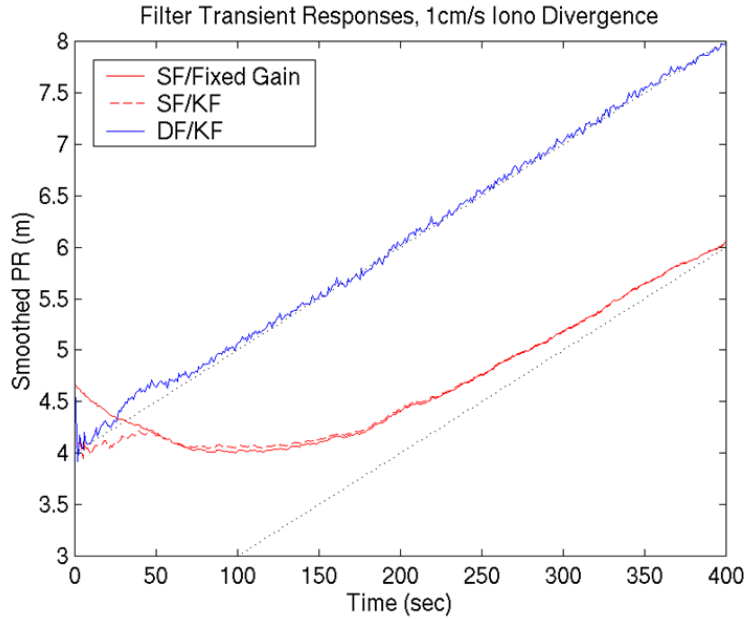


Fig. 21. Simulation of smoothing filter transient responses with ionospheric divergence for steady-state filter time constant of 100 s. The dotted lines show the theoretical steady-state response.

5. Real-time Navigation Processor-based NLOS and Multipath Mitigation

A GNSS navigation processor computes position from measurements of the pseudoranges and sometimes the carrier phase [6][7]. The degree to which these measurements are multipath-contaminated depends on the receiver and antenna design as well as the signal propagation environment. NLOS errors are largely independent of the user equipment design. Where the number of satellites tracked exceeds the minimum required for a navigation solution, the measurements from the receiver can be selected and weighted to minimize the impact of multipath and NLOS errors on the position solution. The C/N_0 and pseudorange rate (Doppler) measurements can assist in this process, as can measurements from signals on different frequencies, where available. Section 5.1 describes methods for weighting each satellite within the position solution using measurements from that satellite alone. Section 5.2 then describes consistency checking, which can identify NLOS reception and large multipath errors by comparing measurements from different satellites.

A navigation filter uses previous as well as current GNSS measurements to compute the position solution, making more information available for detecting and mitigating multipath and NLOS errors. This is particularly useful for mobile applications where the signal propagation environment experienced by the receiver constantly changes. Section 5.3 describes techniques that may be deployed. Aiding information from other navigation and positioning sensors further extends the information available for multipath and NLOS mitigation as discussed in Section 5.4. Note that aiding techniques using 3D mapping are described in Section 7.

5.1 Independent Signal Weighting

All GNSS positioning algorithms that use more than a minimum number of GNSS measurements incorporate a means of weighting those measurements. For example, a single-epoch position solution using m pseudorange measurements, is given by [6]

$$\begin{pmatrix} \hat{\mathbf{r}}_a^+ \\ \hat{\delta}_R^+ \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{r}}_a^- \\ \hat{\delta}_R^- \end{pmatrix} + (\mathbf{H}^T \mathbf{W}_\rho \mathbf{H})^{-1} \mathbf{H}^T \mathbf{W}_\rho \begin{pmatrix} \rho_C^1 - \hat{\rho}_C^{1-} \\ \rho_C^2 - \hat{\rho}_C^{2-} \\ \vdots \\ \rho_C^m - \hat{\rho}_C^{m-} \end{pmatrix} \quad (70)$$

Where \mathbf{r}_a is the Cartesian position of the user antenna with respect to the center of the Earth at time t_a ; δ_R is the receiver clock offset, expressed as a range; \mathbf{H} is the measurement matrix; \mathbf{W}_ρ is the weighting matrix; ρ_C^j is the pseudorange from satellite j to the user, corrected for the estimated satellite clock offset, ionosphere propagation delay and troposphere propagation delay; $\hat{\cdot}$ denotes an estimated value with the superscript $-$ denoting a predicted value from previous measurements and the superscript $+$ denoting a value obtained from the current set of measurements. The user position vector and measurement matrix must be expressed in the same coordinate system, e.g., Earth-Centered-Earth-Fixed (ECEF) or local-level.

The predicted pseudoranges are given by

$$\hat{\rho}_C^{j-} = \sqrt{\left[\Delta \mathbf{C}_E^I \hat{\mathbf{r}}_j^E(t_{st}^j) - \hat{\mathbf{r}}_a^{E-}(t_a) \right]^T \left[\Delta \mathbf{C}_E^I \hat{\mathbf{r}}_j^E(t_{st}^j) - \hat{\mathbf{r}}_a^{E-}(t_a) \right]} + \hat{\delta}_R^- \quad (71)$$

where $\hat{\mathbf{r}}_j^E(t_{st}^j)$ is the Cartesian ECEF position of satellite j at the measured time of signal transmission, t_{st}^j , $\hat{\mathbf{r}}_a^{E-}(t_a)$ is the user ECEF antenna predicted position at the receiver measurement time (common for all satellites). The Sagnac effect is compensated by the rotation matrix, $\Delta \mathbf{C}_E^I$, given by

$$\Delta \mathbf{C}_E^I = \mathbf{C}_E^I(t_a - t_{st}^j) = \begin{pmatrix} \cos \Delta \theta_a^j & \sin \Delta \theta_a^j & 0 \\ -\sin \Delta \theta_a^j & \cos \Delta \theta_a^j & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (72)$$

Where $\Delta \theta_a^j = \omega_{IE}(t_a - t_{st}^j)$ and ω_{IE} is the Earth rotation rate.

The measurement matrix is given by

$$\mathbf{H} = \begin{pmatrix} -e_{a1,x} & -e_{a1,y} & -e_{a1,z} & 1 \\ -e_{a2,x} & -e_{a2,y} & -e_{a2,z} & 1 \\ \vdots & \vdots & \vdots & 1 \\ -e_{am,x} & -e_{am,y} & -e_{am,z} & 1 \end{pmatrix} \quad (73)$$

where $\mathbf{e}_{aj} = (\hat{\mathbf{r}}_j - \hat{\mathbf{r}}_a) / |\hat{\mathbf{r}}_j - \hat{\mathbf{r}}_a|$ is the LOS unit vector from the user antenna to satellite j .

Finally, the weighting matrix is given by

$$\mathbf{W}_\rho = \begin{pmatrix} \sigma_{\rho 1}^{-2} & 0 & \cdots & 0 \\ 0 & \sigma_{\rho 2}^{-2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{\rho m}^{-2} \end{pmatrix} \quad (74)$$

where $\sigma_{\rho j}$ is proportional to the error standard deviation of the j^{th} pseudorange measurement. Thus, measurements with smaller error standard deviations are given higher weighting within the position solution. It is here that measurements may be weighted to mitigate the effects of multipath interference and

NLOS reception. A similar approach may be used with carrier phase and carrier-smoothed code measurements. For a weighting matrix, only the relative values of the components matter, i.e., if all terms are multiplied by a constant, it does not affect the position solution. However, if this weighting matrix is used to estimate solution covariance/uncertainty, as is common in high-integrity positioning applications, then properly characterized LOS standard deviations will be important.

A weighting scheme based on the satellite elevation, θ^j , is sometimes used because lower elevation signals are more susceptible to multipath interference and NLOS reception and also exhibit larger errors in the ionosphere and troposphere propagation models. The Radio Technical Commission for Aeronautics (RTCA) model is [56]

$$\sigma_{\rho_j} = a + b \exp(-\theta^j / \theta_0) \quad (75)$$

For commercial transport aviation applications, the coefficients were empirically-determined to be $a = 0.13$ m, $b = 0.56$ m, and $\theta_0 = 10^\circ = 0.1745$ rad. This model and parameters may not be appropriate for other applications. Such a priori weighting usually gives a more accurate position solution than no weighting, but makes no attempt to directly detect multipath interference or NLOS reception.

Another option for weighting is to use the C/N_0 measurements [57]:

$$\sigma_{\rho_j} = \sqrt{\frac{a}{(c/n_0)^j}} \quad (76)$$

where $c/n_0 = 10^{(C/N_0)/10}$ and a is an empirical application-dependent constant. When a good antenna is used that attenuates LHCP signals, this ensures that multipath and NLOS measurements are generally attenuated, minimizing their impact on the position solution. However, multipath interference has to be very strong to impact the C/N_0 measurement and increases it as often as it reduces it. SNR measurements may be used instead of C/N_0 (according to what the receiver measures), but will need to be rescaled.

A dual-polarization antenna has both a conventional RHCP output, which is generally more sensitive to direct signals and an LHCP output, which is generally more sensitive to reflected signals. Processing these outputs in separate channels within the receiver and differencing the two C/N_0 measurements provides an indication of both NLOS reception and strong multipath interference [19]. In order to exploit signal polarization through measurement weighting, tracking is performed independently on the RHCP and LHCP channels, then the relative strength of the LHCP signal is used as an indication of multipath interference or NLOS reception. RHCP measurements made in the presence of a strong LHCP signal are likely corrupted, so measurements are de-weighted accordingly in the position solution (see the weighting matrix (74)). Weighting is proportional to the difference between the RHCP and LHCP C/N_0 , but must accommodate the variability in signal strength that may arise from the two antenna patterns and signal angle of arrival [58].

Another way of detecting strong multipath interference (but not NLOS reception) is by observing the time variation of the C/N_0 measurements. As described in Section 2, the phase offset of a reflected signal with respect to the direct signal changes, so multipath interference will be varying between constructive and destructive, causing C/N_0 to change. For a moving receiver, these changes will occur relatively quickly. For a static receiver, however, it takes several minutes for the phase offset to complete a cycle [59].

The detection process may be accelerated by using multiple frequencies. As the phase offset is frequency dependent, there will often be constructive interference on one frequency when there is destructive interference on the other, enabling multipath to be detected by comparing the inter-frequency C/N_0 difference with its normal value. With two frequencies, this will only work about half the time. A three-frequency detector will therefore be more robust. A potential triple-frequency detection statistic is [60]:

$$S^j = \sqrt{\left[(C/N_0)_{L1}^j - (C/N_0)_{L2}^j - \Delta C_{12}(\theta^j) \right]^2 + \left[(C/N_0)_{L1}^j - (C/N_0)_{L5}^j - \Delta C_{15}(\theta^j) \right]^2} \quad (77)$$

where $(C/N_0)_{Ln}^j$ is the measured carrier-power-to-noise-density ratio (in dB-Hz) of the signal from satellite j on frequency Ln , and ΔC_{12} and ΔC_{15} are, respectively, the average L1–L2 and L1–L5 inter-frequency C/N_0 differences at elevation θ^j . These averages are determined by calibration in a low-multipath environment. Multipath interference is assumed when S^j exceeds a pre-determined threshold, which is modelled as a function of elevation based on the statistical distribution of the detection statistic in a low-multipath environment.

Multipath interference (but not NLOS reception) also leads to inconsistencies between the different ranging measurements output by the receiver. As the phase offset changes, the pseudorange multipath errors vary more than the corresponding carrier-phase multipath errors. Multipath interference can therefore be detected by looking at the time variation of the difference between the pseudorange and the carrier-phase-derived range, essentially the CMC observable (38) discussed in Section 4. Alternatively, as the phase offset is frequency-dependent, the difference between pseudorange measurements made on different frequencies can be examined. For a moving receiver, both of these parameters will vary quickly as the phase offset of a reflected signal changes as the receiver moves. For a static receiver, it takes several minutes for the phase offset to cycle, so multipath interference takes much longer to detect. It can also be masked by changes in the ionosphere propagation delay. Therefore, in multi-frequency receivers, the following MP observables for satellite j derived from the divergence-free code/carrier combinations (66) can be used:

$$\begin{aligned} MP_{12}^j &= \rho_1^j - \left(\frac{f_1^2 + f_2^2}{f_1^2 - f_2^2} \right) \varphi_1^j + \left(\frac{2f_2^2}{f_1^2 - f_2^2} \right) \varphi_2^j \\ MP_{21}^j &= \rho_2^j - \left(\frac{2f_1^2}{f_1^2 - f_2^2} \right) \varphi_1^j + \left(\frac{f_1^2 + f_2^2}{f_1^2 - f_2^2} \right) \varphi_2^j \\ MP_{51}^j &= \rho_5^j - \left(\frac{2f_1^2}{f_1^2 - f_5^2} \right) \varphi_1^j + \left(\frac{f_1^2 + f_5^2}{f_1^2 - f_5^2} \right) \varphi_5^j \end{aligned} \quad (78)$$

In these MP observables, the ionosphere propagation delay cancels so fluctuations in these parameters over time can be attributed to multipath interference alone [61].

For a large vehicle or monitor station with multiple GNSS antennas at different locations, both multipath interference and NLOS reception can be detected by comparing measurements from these antennas. Inconsistencies in C/N_0 (or SNR) measurements on individual signals between receivers can be an indicator of multipath or of NLOS reception affecting one antenna only (inconsistencies on all signals are likely to be due to interference). Similarly, if the pseudorange difference between antennas is not equal to the distance between them projected onto the receiver-satellite LOS, then NLOS reception or multipath

interference is likely to be present. Differencing across satellites may be needed to eliminate the receiver clock offsets.

Once multipath interference or NLOS reception has been detected, by whatever means, its effects on the position solution must be mitigated. The simplest approach is to exclude the affected measurements from the position solution. However, this does not always improve the accuracy, particularly if the ranging error is relatively small, as excluding a measurement can adversely affect the measurement geometry of the positioning solution. Reducing the weighting of an affected measurement within the position solution is a much more flexible approach. That way, the measurement is used where it is needed to improve the signal geometry, but has little impact on the position solution when the measurement geometry is good without that measurement. For multipath, the weighting can also be adjusted according to the strength of the multipath interference that is detected. The main problem is that C/N_0 -based detection techniques do not provide any indication of the path delay, which is the sole determinant of the NLOS ranging error and a major factor in the ranging error due to multipath interference. By contrast, consistency checking, described in the next subsection, is driven by the ranging errors and thus more relevant to positioning performance.

5.2 Consistency Checking

Consistency checks compare quantities calculated from different combinations of measurements to determine if they are consistent; if they are not, a fault is assumed to be present [6]. NLOS reception and severe multipath interference can thus be detected by comparing measurements of signals from different satellites. If multiple position solutions are computed using combinations of signals from different satellites, those obtained using only combinations of the “clean” direct LOS measurements should be in greater agreement with each other than those that include multipath-contaminated and NLOS measurements. Consistency-checking algorithms therefore identify the set of measurements least affected by multipath interference and NLOS propagation. By excluding or down-weighting the contaminated measurements, a more accurate position solution can usually be obtained. The same consistency checking principle is also used for fault detection in receiver autonomous integrity monitoring (RAIM), as described in Chapter 23.

The simplest GNSS consistency checking approach is sequential testing. The first stage of the sequential testing consistency-checking method is to compute a position solution from all of the pseudorange measurements using Eqn. (70). A vector of residuals is then calculated using

$$\delta \mathbf{z}^+ = \begin{pmatrix} \rho_c^1 - \hat{\rho}_c^{1+} \\ \rho_c^2 - \hat{\rho}_c^{2+} \\ \vdots \\ \rho_c^m - \hat{\rho}_c^{m+} \end{pmatrix} \quad (79)$$

where the pseudoranges computed from the position solution are given by

$$\hat{\rho}_c^{j+} = \sqrt{\left[\Delta \mathbf{C}_E^T \hat{\mathbf{r}}_j^E(t_{st}) - \hat{\mathbf{r}}_a^{E+}(t_a) \right]^T \left[\Delta \mathbf{C}_E^T \hat{\mathbf{r}}_j^E(t_{st}) - \hat{\mathbf{r}}_a^{E+}(t_a) \right]} + \hat{\delta}_R^+ \quad (80)$$

A test statistic based on the sum of the squares of the residuals, $\delta \mathbf{z}^{+T} \delta \mathbf{z}^+$, is then compared with the threshold $\sigma_\rho \sqrt{T_{ch}}$, where σ_ρ is the assumed standard deviation of the pseudorange measurements and T_{ch} is the argument of the chi-square distribution corresponding to the required upper confidence limit [62]. The number of degrees of freedom is equal to the number of satellites from which pseudorange

measurements are used, minus the number of parameters in the navigation solution (four in the case of Eqn. (70)). Where the test statistic falls within the threshold, the position solution is accepted. Otherwise, it is assumed that at least one measurement is NLOS, severely multipath-contaminated or subject to another source of error. The measurement with the largest residual is then eliminated as it is least consistent with the others and the process repeats.

This “top down” approach is effective for relatively benign environments where only one or two measurements are contaminated. However, it does not work for dense urban environments where the majority of signals may be NLOS or affected by severe multipath interference. In such cases, the residuals produced from a weighted least-squares solution can be poor indicators of the quality of the individual signals. Measurements from clean signals are as likely to have large residuals as those from contaminated signals, so the wrong measurements may be eliminated in a sequential testing algorithm [63].

An alternative “bottom up” approach to consistency checking is the subset comparison method [63]. A minimal sample set (MSS) is a subset consisting only of the minimum number of GNSS measurements required to produce an exact solution, i.e., four plus any unknown inter-constellation timing biases. The subset comparison method works by generating a series of MSSs and scoring each one according to its consistency with the other measurements. A solution is then computed using the best-scoring MSS plus those other measurements found to be consistent with it. Unlike the sequential testing method, many different measurement combinations are assessed, so the chances of selecting the most consistent set are much higher.

For each MSS, a position solution is computed using those measurements only. If the i^{th} MSS contains the measurements $\mathbf{z}^i \in \mathbf{z}$, where $\mathbf{z} = (\rho_c^1 \quad \rho_c^2 \quad \cdots \quad \rho_c^m)^T$, the position and receiver clock offset is then

$$\begin{pmatrix} \hat{\mathbf{r}}_a^{+i} \\ \hat{\delta}_R^{+i} \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{r}}_a^- \\ \hat{\delta}_R^- \end{pmatrix} + (\mathbf{H}^i)^{-1} (\mathbf{z}^i - \hat{\mathbf{z}}^{i-}) \quad (81)$$

where \mathbf{H}^i comprises the rows of the measurement matrix, \mathbf{H} , given by (73), corresponding to MSS i and the vector of predicted pseudoranges, $\hat{\mathbf{z}}^{i-}$, is determined using (71). Because the selection of each MSS is defined to have the minimum number of rows required to estimate the states, \mathbf{H}^i is square, and will be invertible provided satellite geometry is good enough that it is not numerically singular.

Each MSS is scored according to its consistency with the remaining GNSS measurements. The first step is to calculate a set of residuals for the full set of measurements, $\delta \mathbf{z}^{+i}$. This is calculated using (79) and (80) with $\hat{\mathbf{r}}_a^{+i}$ and $\hat{\delta} b_c^{+i}$ substituted for $\hat{\mathbf{r}}_a^+$ and $\hat{\delta} b_c^+$. Note that the residuals for those measurements within the MSS will then be zero. Those measurements outside the MSS which have residuals with magnitudes less than the empirically-determined threshold δz_{\max} are considered part of that MSS’s consensus set (CS). The MSS and its associated CS are then scored using the cost function

$$C^i(\mathbf{e}^i) = \sum_{j=1}^m k(\delta z_j^{+i}, \delta z_{\max}) \quad (82)$$

where k is the individual measurement cost function. Better performance is obtained if this is weighted in the same way as a weighted least squares solution. Thus,

$$k(\delta z_j^{+i}, \delta z_{\max}) = \begin{cases} |\delta z_j^{+i}| / \sigma_{\rho j} & |\delta z_j^{+i}| \leq \delta z_{\max} \\ \delta z_{\max} / \sigma_{\rho j} & |\delta z_j^{+i}| > \delta z_{\max} \end{cases} \quad (83)$$

where $\sigma_{\rho j}$ is proportional to the error standard deviation of the j^{th} pseudorange measurement; the weighting schemes described in Section 5.1 may be used.

The MSS with the lowest cost function, C^i , and its associated CS then form the set of measurements, \mathbf{z}^f , used to compute the final position and timings solution using

$$\begin{pmatrix} \hat{\mathbf{r}}_a^+ \\ \hat{\delta}_R^+ \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{r}}_a^- \\ \hat{\delta}_R^- \end{pmatrix} + \left((\mathbf{H}^f)^T \mathbf{W}_\rho^f \mathbf{H}^f \right)^{-1} (\mathbf{H}^f)^T \mathbf{W}_\rho^f (\mathbf{z}^f - \hat{\mathbf{z}}^{f-}) \quad (84)$$

where \mathbf{H}^f are the rows of the measurement matrix, \mathbf{W}_ρ^f the rows and columns of the weighting matrix, and $\hat{\mathbf{z}}^{f-}$ the set of predicted pseudoranges corresponding to the final measurement set. They are computed using (73), (74) and (71), respectively.

It is not necessary to test all possible MSSs as there will typically be more than one combination of MSS and CS that form the final measurement selection. Instead, the random sample consensus (RANSAC) technique may be used [64]. This randomly generates MSSs until a sufficient number have been generated for the probability that none of the MSSs are outlier free has dropped below a predetermined significance level [63]. Another approach based on applying chi-square tests to different measurement combinations is described in [65].

5.3 Using a Filtered Navigation Solution

As discussed in Section 4, carrier smoothing of code pseudoranges, can provide substantial attenuation of code multipath. Carrier smoothing is performed separately for measurements from each satellite, producing a set of smoothed pseudorange measurements which are then processed in the same way as raw pseudoranges. Thus, signal weighting and consistency checking can be applied as described in Sections 5.1 and 5.2, respectively. Therefore, while carrier smoothing does not directly mitigate NLOS reception errors—which affect the code and carrier measurements in the same way—the reduction in pseudorange noise does aid in detecting NLOS reception.

Instead of filtering measurements in the range domain, separately for each satellite, filtering can also be accomplished in the position domain, as part of the navigation solution. This is done by replacing the weighted least squares navigation solution in (70) with an extended Kalman filter (EKF) that maintains a continuous position, velocity and time (PVT) solution with an associated error covariance. This is predicted forward in time, using the velocity solution to predict the change in position, and the receiver clock drift to predict the change in clock offset. New GNSS pseudorange and pseudorange rate (Doppler) measurements are then used to correct the predicted PVT solution, weighted according to the relative error covariance of the measurements and the predicted solution. Further details are presented in [6] and in Chapter 46.

Independent signal weighting to reduce the impact of multipath interference and NLOS reception can be implemented in an EKF using a similar approach as that described in Section 5.1. However, there are some differences. A measurement noise covariance matrix is used instead of a weighting matrix (noting that one is the inverse of the other). This represents only errors that vary rapidly with time, including multipath and NLOS reception where the receiver is moving. Unlike a weighting matrix, a measurement noise covariance

matrix must also be correctly scaled to ensure optimal weighting of old and new information within the filter [6]. The other main difference is that there are carrier-derived pseudorange rate measurements to weight as well as the pseudoranges; carrier-phase-derived measurements may also be used, where available. Weighting should be performed under the assumption that all measurements from a given satellite are affected when multipath interference and/or NLOS reception is present.

Consistency checking can also be performed as described in Section 5.2, with those measurements that pass the consistency-checking process then input to the EKF. All measurements from the same satellite should be accepted or rejected together as propagation paths for different frequency components are highly correlated and are likely also corrupted, even if the errors have not yet been detected. However, an EKF also enables measurements to be compared for consistency with the navigation solution predicted from previous epochs. This is known as innovation filtering. The measurement innovation vector of an EKF is [6]

$$\delta \mathbf{z}_k^- = \mathbf{z}_k - \mathbf{h}(\hat{\mathbf{x}}_k^-) \quad (85)$$

where \mathbf{z}_k is the set of measurements at epoch k , $\hat{\mathbf{x}}_k^-$ is the set of state estimates (typically position, velocity and time) at epoch k , predicted forward from the previous epoch, and \mathbf{h} is a nonlinear measurement function that expresses the measurements as a function of the states. The measurement vector will typically comprise pseudoranges and pseudorange rates (often in the form of carrier phase delta ranges).

The covariance of the innovations, $\mathbf{C}_{\delta \mathbf{z},k}^-$, is a step in the computation of the Kalman Filter gain, and comprises the sum of the measurement noise covariance and the error covariance of the state estimates transformed into measurement space:

$$\mathbf{C}_{\delta \mathbf{z},k}^- = \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k \quad (86)$$

where \mathbf{P}_k^- is the error covariance matrix of the predicted state estimates, \mathbf{H}_k is the measurement matrix, comprising the Jacobian matrix of partial derivatives of each measurement with respect to each state, and \mathbf{R}_k is the measurement noise covariance matrix. The normalized innovation for the i^{th} scalar measurement is given by

$$\delta \bar{z}_{k,i}^- = \frac{\delta z_{k,i}^-}{\sqrt{\mathbf{C}_{\delta z,k,i,i}^-}} \quad (87)$$

Outlying measurements can then be detected by comparing each $\delta \bar{z}_{k,i}^-$ with a threshold. A higher threshold minimizes the false alarm rate while a lower threshold minimizes the missed detection rate. If two thresholds are used, measurements above the higher threshold can be rejected while measurements falling between the two thresholds are de-weighted. Test statistics can also be computed from a sequence of normalized measurement innovations, increasing sensitivity at the expense of response time [6]. Examining a sequence of innovations also enables NLOS reception and multipath interference to be distinguished, with the former indicated by a bias and the latter by a larger variance than normal [66].

Vector tracking combines PVT estimation and GNSS signal tracking into a single estimation algorithm. The navigation filter inputs code and carrier discriminator measurements instead of ranging measurements and its PVT solution is used to generate the numerically controlled oscillator (NCO) commands within the receiver that control the reference code and carrier generation. Full details are presented in Chapter 16. In

vector tracking, low- C/N_0 discriminator measurements are automatically de-weighted so NLOS measurements have less impact on the position solution [67]. Innovation filtering can also be performed on the discriminator measurements to enable rejection or de-weighting of NLOS and strongly multipath-contaminated signals [68].

5.4 Using Aiding Information

A dead-reckoning navigation system measures motion using inertial sensors, wheel-speed sensors, Doppler radar, Doppler sonar, or another technology, and then integrates that motion to update its position solution. As the motion measurement errors are also integrated, the position accuracy degrades with time. Consequently, dead-reckoning technologies are normally integrated with position-fixing technologies, such as GNSS, typically using Kalman filter-based estimation. Details are presented in Chapter 46 and in [6][69][70]. In such an integrated navigation system, the motion between GNSS epochs is measured instead of predicted. Consequently, in innovation filtering (Section 5.3), the predicted navigation solution used to compute the GNSS measurement innovations is more accurate, making the innovation filtering more sensitive and thus making NLOS reception and multipath interference easier to detect.

As described in [6], GNSS can also be integrated with other radio positioning technologies, environmental feature matching systems (e.g. magnetic anomaly matching), map matching and/or terrain height aiding (see Section 7). These all increase the amount of information available to the navigation filter, boosting the sensitivity of innovation filtering. Chapter 35-43 in this book are dedicated to techniques that use radio signals of opportunities for navigation, while Chapters 44-52 describe non-radio-based navigation technologies.

A sky-pointing camera with a panoramic lens or an array of cameras can produce an image of the entire field of view above the receiver's masking angle. Where the orientation of the camera is known (requiring integration with inertial sensors or multi-antenna interferometric GNSS), the blocked lines of sight may be determined from the image. By comparing these with the satellite azimuths and elevations, NLOS signals can be identified and excluded from the PVT solution [71][72].

6. Post-processing Techniques for Multipath Mitigation

Some methods for removing multipath effects rely on post-processing of collected GNSS data. Post-processing approaches can use batch estimation techniques; exploit constellation repeat cycles; leverage additional information, such as known antenna motion; and employ computationally-intensive methods, such as electromagnetic (EM) ray tracing.

6.1 Fixed or Repeatable Geometry

For a stationary receiver, installed in an essentially static environment, the repeating satellite geometry of each GNSS constellation produces repeating multipath effects that can be identified and removed. This repeat time varies by constellation: the GPS constellation repeats every sidereal day, GLONASS every eight, Galileo every ten [73], and BeiDou every seven [74]. Techniques exploiting repeating satellite geometry are referred to as sidereal filtering, in reference to the constellation repeat time of GPS for which the technique was first conceived [75]. Multipath corrections may be applied in the position domain or the observation domain. In Position Domain Sidereal Filtering (PDSF), a series of position residuals (e.g., East/North/Height errors) from one day are subtracted from a series of corresponding position states from a subsequent day [76]. The position residuals from the first day may be low-pass filtered to avoid amplifying

noise. In the post-processing case considered here, the input to the position estimation for each day should be checked to ensure only satellites visible both days are used in case the constellation changes (e.g., a satellite outage) [77].

Observation Domain Sidereal Filtering (ODSF) involves correction of the range measurements themselves: measurement residuals from one day are low-pass filtered and subtracted, per satellite, from a subsequent day's series of corresponding measurements. The low-pass filter bandwidth must be sufficiently less than the multipath fading frequency (see Eqns. (4) and (5)).

Greater accuracy in multipath removal can be achieved by using more sophisticated alignment techniques. Although the GPS constellation nominally repeats every 86164 seconds, for example, Modified Sidereal Filtering (MSF) uses a repeat time of 86155 seconds, accounting for the westward drift of the satellite planes due to the oblateness of the Earth [79]. Orbit periods of GPS satellites vary by about eight seconds over a year, however, and spacecraft maneuvers can change orbit periods by more than 100 seconds [77]; the repeat of a satellite ground track is a better indicator of repeating multipath than orbit period. The Aspect Repeat Time (ART) method takes this approach by finding the time shift that maximizes the dot product between the two user-to-satellite LOS vectors [78],[80]. Note that in each of these methods, accurate alignment of the previous day's residuals with the current observations relies on a sufficiently high rate data (on the order of 1 Hz or greater). Noise in the multipath correction may be reduced by averaging residuals from multiple days [81].

Another approach that exploits repeatable geometry is to map measurement residuals onto a sky plot by azimuth and elevation, thus forming a hemispherical template of multipath corrections [82],[83]. In most cases, however, accounting for high frequency multipath effects would require a prohibitively fine discretization of the hemisphere and generating a complete template can take months or years [76]. This technique is still useful for spacecraft applications, where reflectors are very close to the receive antenna, implying lower frequency errors (see Eqns. (4) and (5)), and the environment may be indefinitely static.

6.2 Estimation of Reflected Ray Properties

Estimation of the multipath parameters (α_i , Δ_i , and ϕ_i for each of $i = 1, \dots, n$ reflected signals in (6)) allows for reconstruction of the signals themselves, thereby enabling correction of multipath errors via any of the previously discussed methods. While methods discussed in Section 3.3, such as MEDLL, perform estimation in real-time, others benefit from a post-processing approach. Due to the impossibility of perfect isolation of direct and reflected signals in the cases of interest (i.e., cases with multipath-induced errors), one approach is to rely on differences in the evolution of direct and reflected signals in a time series of data. Another approach is to thoroughly model the receiver environment and estimate the reflected signal properties through EM ray tracing.

Multipath error is highly sensitive to antenna motion—moving an antenna just half a wavelength will significantly change the relative phases of the direct and reflected signals. As an antenna is moved, projection of the antenna motion onto the different signal arrival directions varies. By generalizing the multipath model used in MEDLL to include antenna motion, a known antenna motion time history can be used to estimate the multipath parameters of each reflected signal from the in-phase and quadrature accumulations [84]. This has been demonstrated experimentally with some success (e.g., reduction of overall accumulated delta pseudorange RMS from 10 to 6 m) but is still in development.

Rather than estimating multipath parameters from measurements, these parameters can be calculated from a known receiver environment and the laws that govern the propagation of EM radiation. Representing a propagating EM field as a ray is a simplification known as geometrical optics (GO) [85]. As long as object dimensions are much larger than a wavelength, the interaction of this ray with surfaces and different mediums is described by the GO laws of reflection and refraction—though these do not apply to edges. The Uniform Geometrical Theory of Diffraction (UTD) expands the GO laws to handle complex bodies by introducing diffraction to describe the scattering of rays at edges [86][87]. When electrical properties of an environment are fully defined, EM ray tracing can be used to describe the path taken by an EM ray, including cases of multipath.

In a typical approach to calculating multipath through EM radiation modeling, structures in the receiver environment are first decomposed into simple geometric shapes. These shapes are assigned diffraction and reflection coefficients according to their material properties and the canonical shapes characterized in the UTD. All field components contributing to the field at the antenna phase center are traced through the environment and finally the individual field components are summed together. For example, [92] describes an individual electric field component at point \mathbf{r}' :

$$\mathbf{E}^{r,d}(\mathbf{r}') = \mathbf{E}^i(\mathbf{r}) \mathbf{D}^{r,d} \mathbf{A}^{r,d}(s) e^{-jks} \quad (88)$$

where the superscript r,d indicates that the field may arise from either interaction (reflection or diffraction), $\mathbf{E}^i(\mathbf{r})$ is the incident field arriving at \mathbf{r} from the source or the previous interaction point, $\mathbf{D}^{r,d}$ is a complex reflection or diffraction coefficient, $\mathbf{A}^{r,d}(s)$ a spreading factor due to distance s , and e^{-jks} a phase term with wavenumber k . The total field at the receive antenna is the sum of the LOS field, reflected fields, and diffracted fields.

Ray tracing software is used to perform the computationally intensive task of tracing all significant field components through the receiver environment and applying interaction coefficients. Software must be selected according to a number of factors, such as ease of use, visualization capabilities, speed, and accuracy. Most software tools use UTD methods [89],[90],[91], but accuracy is highly dependent on the accuracy of structure models [92]. Ray tracing faces significant challenges in terrestrial applications, but has found use in simple or well understood environments, such as spacecraft [93]. Correcting code multipath errors requires knowledge of reflected signal path delay to within one tenth of a wavelength (~ 2 cm), and therefore centimeter-level knowledge of antenna position, phase center, and environment features is needed. This is further complicated by imperfect phase reversal upon reflection (most buildings or other structures are not flat over the Fresnel zone at GNSS wavelengths). As discussed in Section 7, however, it is practical to use ray tracing for NLOS correction, as much greater modeling errors can be tolerated. Similarly, ray tracing can be used to provide rough estimates of multipath relative amplitude, α_i , which in turn can be used to determine pseudorange error weighting for the position solution as in (74).

6.3 Multipath Characterization

Characterization of the multipath environment is important for antenna placement and measurement weighting. But it can be difficult to identify multipath-induced errors from among the many error sources. As described in Section 4.1, pseudorange multipath errors can be isolated by processing CMC data. These techniques exploit the fact that pseudorange and carrier phase share most signal-in-space error sources, but

have different multipath errors. Consider code and carrier measurements in (31). The single-frequency "code-minus-carrier" multipath observable (38) is the difference between these two range measurements,

$$\begin{aligned}\chi_L &= \rho_L - \varphi_L = 2I_L + \delta\rho_{ML} - \delta\varphi_{ML} + \varepsilon_{\rho L} - \varepsilon_{\varphi L} - N_L\lambda_L \\ &\approx 2I_L + \delta\rho_{ML} + \varepsilon_{\rho L} - N_L\lambda_L\end{aligned}\quad (89)$$

in which the geometric range, tropospheric delay, and clock terms cancel [94][95]. Carrier multipath error and carrier noise terms are negligible relative to the code terms and can be ignored. The integer carrier ambiguity can be removed by subtracting the mean of χ_L , but note that this removes any bias associated with the multipath errors. Finally, the ionosphere terms must be removed. This can be achieved by estimating and removing slow trending in the measurement (on the order of hours) [95], or by producing an ionosphere free combination of the code and carrier range measurements with multiple frequencies—see Section 4.3.3 and Equation (78). The resulting measure of pseudorange multipath errors can be employed to de-weight or ignore corrupted measurements [96].

Carrier phase multipath errors can be isolated with differential phase techniques (i.e., the difference in phase measured by two antennas) by considering the relationship between measured SNR and carrier phase error. If known factors affecting gain are removed (e.g., first order transmitter and receiver motion), variation in the adjusted SNR can be attributed to multipath and used to estimate and correct phase errors [97]. The expression for carrier phase error for a single reflected signal stated in Equation (16) can be expanded to an arbitrary number of signals and, for small range lag ($|x - \delta| < d/2$) and small α_i , approximated as

$$\delta\varphi_M \approx \frac{\sum_{i=0}^n \alpha_i \sin(\phi_i)}{1 + \sum_{i=0}^n \alpha_i \cos(\phi_i)}\quad (90)$$

The amplitude of the combined signal, for direct signal amplitude A_d and small α_i , can be approximated as

$$A_c \approx A_d \sum_{i=0}^n \alpha_i \cos(\phi_i).\quad (91)$$

This multipath-induced variation in the amplitude of the combined signal is evident in the adjusted SNR and can be used to estimate the relative amplitude and phase of each multipath component. The sum of these components forms a profile for removing carrier phase multipath errors [98].

Assuming the multipath frequency is constant over the time interval considered,

$$\phi_i = \omega_i t + \phi_{0,i},\quad (92)$$

where ω_i is the angular frequency of the i^{th} multipath component and $\phi_{0,i}$ is the initial phase offset, and the number of multipath components and their frequencies can be identified through spectral analysis (e.g., fast Fourier transform methods). An example is shown in Fig. 22, in which the signal power exhibits a multipath-induced oscillation, and a single multipath component with frequency 0.25 Hz can be identified in the frequency domain, despite the constant frequency assumption in (92) only holding approximately.

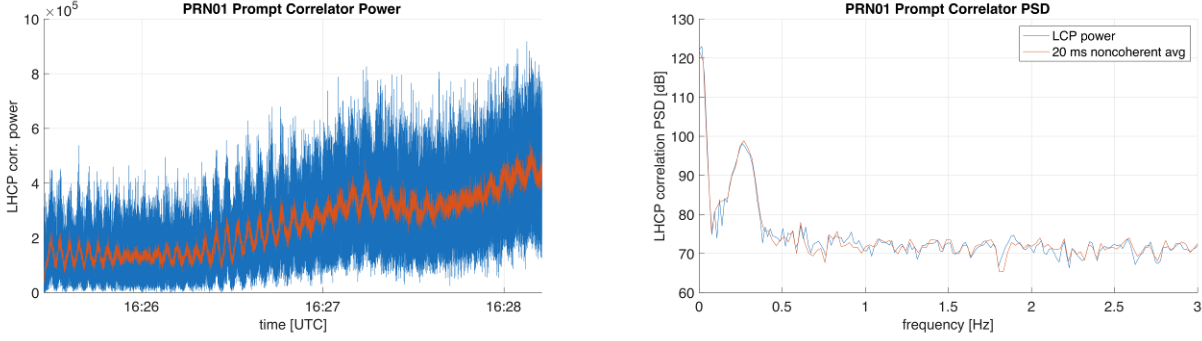


Fig. 22 Prompt correlator power (left) and power spectral density (right) of PRN 1 measured during Hubble Servicing Mission 4 [20].

The amplitudes and initial phases of each multipath component are estimated through a least-squares fit to the model:

$$\begin{bmatrix} \cos(\omega_1 t_k) & -\sin(\omega_1 t_k) & \cdots & \cos(\omega_n t_k) & -\sin(\omega_n t_k) \\ \vdots & \vdots & & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \alpha_1 A_d \cos(\phi_{0,1}) \\ \alpha_1 A_d \sin(\phi_{0,1}) \\ \vdots \\ \alpha_n A_d \cos(\phi_{0,n}) \\ \alpha_n A_d \sin(\phi_{0,n}) \end{bmatrix} = \begin{bmatrix} SNR_M(t_k) \\ \vdots \end{bmatrix} \quad (93)$$

where the multipath SNR time series, SNR_M , is computed by subtracting the estimated direct SNR (calculated according to a link budget and known geometry) from the measured and adjusted SNR [98]. Note that the real-time SNR estimate must not be too heavily filtered in order to observe changes in the total SNR. The resulting ω_i , $\phi_{0,i}$, and α_i estimates are used to calculate the multipath frequency via (92), and in turn the carrier phase error in (90). The spectral analysis used to determine ω_i does not determine sign, however, so measurements from two antennas are required. An additional least squares fit must be performed according to the model

$$\begin{bmatrix} \delta\varphi_{M,1,1}(t_k) & \cdots & \delta\varphi_{M,n,1} & \cdots & \delta\varphi_{M,1,2} & \cdots & \delta\varphi_{M,n,2} \\ \vdots & & \vdots & & \vdots & & \vdots \end{bmatrix} \begin{bmatrix} s_{1,1} \\ \vdots \\ s_{n,1} \\ s_{1,2} \\ \vdots \\ s_{n,2} \end{bmatrix} = \begin{bmatrix} DPHS_{resid}(t_k) \\ \vdots \end{bmatrix} \quad (94)$$

where $\delta\varphi_{M,i,j}$ is the estimated carrier phase error for multipath component i and antenna j , $DPHS_{resid}$ the motion-corrected difference between the phase measured at each of the two antennas, and $s_{i,j}$ the carrier phase error sign [98]. After these signs are computed, a differential phase correction profile can be constructed and applied to the differential phase data to produce carrier multipath free differential phase measurements.

With the availability of triple-frequency GNSS signals, geometry/ionosphere-free phase combinations can be formed [99]. These are useful for characterizing overall carrier phase multipath error levels, but not the

multipath error on a single frequency. Using the measurement models in (31), the triple-frequency phase combination suggested in [99] can be expanded to obtain:

$$\begin{aligned}\delta\varphi_{M123} &= \lambda_3^2(\varphi_1 - \varphi_2) + \lambda_2^2(\varphi_3 - \varphi_1) + \lambda_1^2(\varphi_2 - \varphi_3) \\ &= (\lambda_3^2 - \lambda_2^2)\delta\varphi_{M1} + (\lambda_1^2 - \lambda_3^2)\delta\varphi_{M2} + (\lambda_2^2 - \lambda_1^2)\delta\varphi_{M3} + (\text{Noise}) + (\text{Ambiguities})\end{aligned}\quad (95)$$

where the subscripts 1, 2, and 3 denote the frequencies. This can be done with a single antenna/receiver and can be used to estimate overall multipath error statistics, assessing error budgets and comparing different antenna sites.

7. 3D Mapping-Aided GNSS

GNSS position accuracy is degraded in dense urban environments because buildings block and reflect signals. 3D mapping of buildings (together with knowledge of satellite positions) enables prediction of which signals are affected where. By using this information in GNSS positioning algorithms, position error may be reduced from tens of meters to a few meters [100]. Techniques may be divided into terrain height aiding, 3D mapping-aided (3DMA) GNSS ranging, and shadow matching. These are described in turn, followed by a discussion of system implementation issues.

7.1 Terrain Height Aiding

For most land positioning applications, the GNSS receive antenna may be assumed to be at a known height above the terrain. By using a digital terrain model (DTM), also known as a digital elevation model (DEM), the position solution may be constrained to a surface. Terrain height aiding was used in the early days of GPS to enable positioning with a limited number of satellites. By effectively removing a dimension from the position solution, the accuracy of the remaining dimensions is improved. In an open environment, terrain height aiding only significantly improves the vertical positioning and timing accuracies. However, where the signal geometry is poor, such as in dense urban areas, horizontal accuracy can be improved by nearly a factor of two [101].

Terrain height aiding is incorporated into a conventional least-squares or EKF positioning algorithm by adding a virtual ranging measurement [6][101][102]. This comprises the distance, r_{ea} , from the center of the Earth to terrain at the predicted horizontal position, adjusted for any known vertical displacement of the user antenna from the terrain. The DTM will provide the terrain height at a series of grid points, so interpolation is necessary. Bicubic interpolation using the nearest four grid points is effective [101]. Clearly the more accurate the predicted horizontal position is, the more accurate the terrain height will be. Therefore, the positioning algorithm should be iterated several times, using the horizontal position solution from the previous iteration to compute the terrain height at each iteration.

The positioning algorithm described in Section 5.1 can be augmented with terrain height aiding by adding an additional measurement, $\tilde{r}_{a,T}^E$, to Equation (70). The terrain height aiding measurement innovation, $\tilde{r}_{a,T}^E - |\hat{\mathbf{r}}_a^-|$ is added to the measurement residual vector of (70). The measurement matrix, (73), is augmented with the additional row

$$\mathbf{H}_{m+1} = \begin{pmatrix} e_{a,x} & e_{a,y} & e_{a,z} & 0 \end{pmatrix}\quad (96)$$

where $\mathbf{e}_a^E \approx \hat{\mathbf{r}}_a^{E-} / |\hat{\mathbf{r}}_a^{E-}|$ in ECEF coordinates or $\mathbf{e}_a^L \approx (0 \ 0 \ 1)^T$ in local level coordinates. An additional row and column is also added to the weighting matrix, \mathbf{W}_ρ , to weight the height-aiding measurement within the position solution.

Terrain height aiding may also be used to improve the robustness of consistency checking (Section 5.2) [63][103].

7.2 3D Mapping-Aided Ranging

Using ray tracing or projection techniques, 3D maps of cities can be used to predict which signals are blocked by buildings, and thus are NLOS when received, and which are directly visible. Direct LOS signals may or may not be subject to multipath contamination. This process can be accelerated by using pre-computed building boundaries. The building boundary at a given position comprises the elevation threshold below which satellite signals are blocked for each azimuth. Satellite visibility can then be predicted very quickly simply by comparing the satellite elevation with the building boundary elevation at the appropriate azimuth. These satellite visibility predictions can be used to aid ranging-based positioning in a number of different ways. Where the position is already known to within a few meters, it is possible to predict which signals are NLOS with reasonable accuracy and simply exclude them from the position solution (assuming there are sufficient direct LOS signals) [104][105]. Otherwise, which signals are directly visible depends on the true position, which is not known. A simple approach is to determine the proportion of candidate positions at which each signal is predicted to be directly visible and use this to weight each measurement within the position solution and to aid consistency checking [101]. This approach typically improves the positioning accuracy by 20-25% and can handle initialization errors of about 100 m.

To make the best use of satellite visibility prediction, a conventional least-squares (or EKF) positioning algorithm should be replaced by an algorithm that scores candidate position hypotheses according to the difference between the measured and predicted pseudoranges, assuming LOS propagation. The receiver clock offset and any inter-constellation timing biases may be eliminated by differencing measurements across satellites. Different assumptions about the error distribution can then be made at different candidate positions according to which signals are predicted to be LOS or NLOS at each position. Thus, a symmetric error distribution can be assumed for LOS signals and an asymmetric distribution for NLOS signals with the scoring adjusted accordingly [106]. The candidate positions may be distributed in a regular grid or semi-randomly (like in a particle filter). The search area containing those position candidates is centered at either the conventional GNSS position solution or a position predicted forward from previous epochs. The size of the search area is then based on the uncertainty of the initializing position.

Ray tracing enables the path delay of a reflected GNSS signal to be predicted as discussed in Section 6.2. NLOS reception errors may then be corrected, enabling NLOS signals to contribute to an accurate position solution. However, accurate correction of NLOS errors requires an accurate position solution. If the position is already known to within a few meters, alternate computation of the position solution and NLOS corrections may be iterated until they converge. For larger uncertainties, multiple starting positions will be needed to ensure convergence. A more powerful approach adds NLOS error prediction to positioning by scoring candidate position hypotheses. Appropriate NLOS corrections are then computed for each candidate position. Both grid-based and particle-based methods have demonstrated positioning accuracies within 2 meters [107][108][109].

7.3 Shadow Matching

Shadow matching is a complementary GNSS positioning technique that determines position by comparing predicted and measured C/N_0 . Unlike conventional GNSS positioning, it does not use the ranging measurements, but rather, is akin to “RF fingerprinting” techniques used for indoor location systems. Using 3D mapping and the satellite positions, each GNSS signal is predicted to be directly visible in some areas and blocked, e.g., shadowed in other areas. Shadow matching therefore assumes that the user is in one of the directly visible areas if the received SNR is high and in one of the shadowed areas if the SNR is low or the signal is not received at all. Figure 23 illustrates the general principle. Repeating this for each GNSS signal enables the area within which the user may be found to be reduced [110].

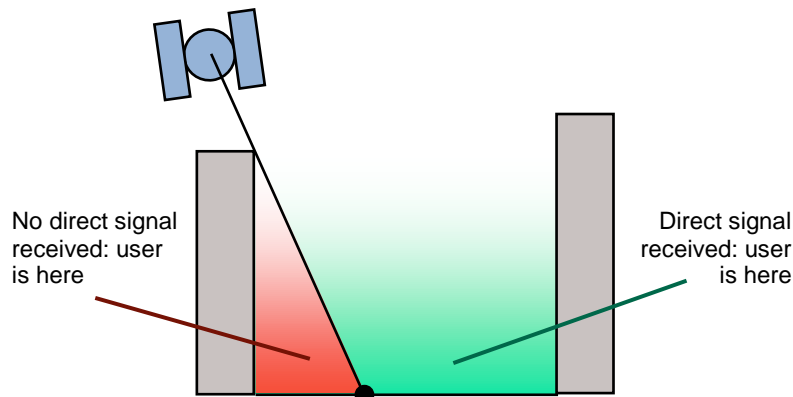


Fig. 23: Principle of Shadow Matching

In practice, there can be overlaps in the SNR distributions of direct LOS and NLOS signals, particularly when a smartphone antenna is used. Furthermore, real urban environments and signal propagation behavior are more complex than it is possible to represent using 3D mapping [111]. Therefore, a practical shadow-matching algorithm works by scoring a grid of candidate positions according to the degree of correspondence between the satellite visibility predictions and the SNR measurements. This enables inaccuracies in the process to be treated as noise, so a correct position is still obtained provided there is sufficient “signal”.

Figure 24 shows the stages of a typical shadow-matching algorithm [112]. Before using shadow matching, the context must be determined [113]. Shadow matching requires an outdoor urban environment as it does not work indoors and is not needed in an open environment where conventional GNSS positioning works well. The first step is to establish the search area using an approximate position (e.g., from conventional GNSS) and an associated uncertainty. A search radius of a few tens of meters is typically needed. Within this search area, a grid of candidate positions is established; indoor locations may be omitted. Shadow matching works well with a grid spacing of 1 m, but can operate with a larger spacing.

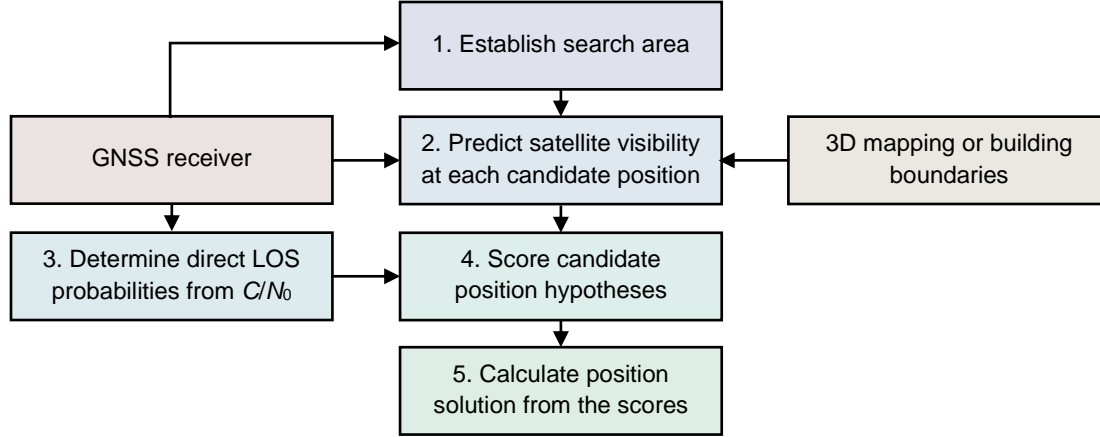


Fig. 24: Stages of a Typical Shadow-Matching Algorithm

Next, 3D mapping is used to predict the satellite visibility at each position, either directly or via pre-computed building boundaries. A 3D city model is necessarily an approximation of the true environment, and a ray is only an approximation of a GNSS signal, the Fresnel radius of which can exceed a meter in urban environments. Thus, although the satellite visibility predictions are Boolean, it is better to treat the probability that a signal is direct LOS as non-Boolean. For example, experimental tests have shown that a LOS probability of 0.8 can be assumed if the signal is predicted from the 3D mapping to be LOS and a LOS probability of 0.2 can be assumed for signals predicted to be NLOS [112].

The third step is to determine which of the received signals are direct LOS from the C/N_0 (or SNR) measurements output by the receiver. Clearly, if the C/N_0 is close to nominal, the signal is likely to be direct LOS. Conversely, if the C/N_0 is just above the code tracking threshold or no signal is tracked, then the direct signal path is almost certainly blocked. However, intermediate values of C/N_0 can be more difficult to classify as some NLOS signals can be very strong while direct signals can be attenuated by people and foliage. As discussed previously, mobile phone antennas are linearly polarized, and present a particular problem as they do not distinguish between RHCP and LHCP signals, and the gain pattern experiences a null along the axis of the antenna. Therefore, an empirically-determined function that expresses the direct LOS probability as a function of C/N_0 should be used. This function can be derived from C/N_0 data collected at known locations [112]. Because of their different antenna characteristics, different models are needed for professional-grade, consumer-grade and mobile phone GNSS user equipment.

The fourth stage is to score each of the candidate positions. The probability that the predicted and measured satellite visibility match is

$$\begin{aligned}
 P_{ij} &= p(\text{LOS} | C/N_0)_j p(\text{LOS} | \text{map})_{ij} + (1 - p(\text{LOS} | C/N_0)_j) (1 - p(\text{LOS} | \text{map})_{ij}) \\
 &= 1 - p(\text{LOS} | C/N_0)_j - p(\text{LOS} | \text{map})_{ij} + 2p(\text{LOS} | C/N_0)_j p(\text{LOS} | \text{map})_{ij}
 \end{aligned} \tag{97}$$

where $p(\text{LOS} | \text{map})$ is the predicted direct LOS probability, $p(\text{LOS} | C/N_0)$ is the observed direct LOS probability, j is the satellite and i is the candidate position. A score for each candidate position can then be determined by multiplying the match probabilities for each signal. Thus,

$$\Lambda_i = \prod_j P_{ij} \tag{98}$$

The output of the scoring process is thus a likelihood surface, giving the likelihood of each candidate position. This can be converted to a probability density function by normalizing it so that it integrates to unity.

The final step in the shadow-matching process is to compute the position solution. A simple approach is to take a weighted average of the candidate positions. Thus:

$$\hat{\mathbf{p}} = \frac{\sum_i \Lambda_i \mathbf{p}_i}{\sum_i \Lambda_i} \quad (99)$$

where \mathbf{p}_i is the position of the i^{th} candidate, which may be expressed as Cartesian, curvilinear, or projected coordinates. However, the likelihood surface can sometimes be multimodal, resulting in a position solution that is the average of several possibilities. This can be accounted for by increasing the position uncertainty in these cases. More sophisticated approaches include extracting multiple position hypotheses from the likelihood surface and putting the full likelihood surface into an integration filter. Several groups have demonstrated multi-epoch shadow matching using a particle filter, achieving cross-street positioning accuracies of better than 3m [114][115][116][117].

7.4 System Implementation

Terrain height aiding, 3DMA GNSS ranging, and shadow matching should not be thought of as competitors. Best performance is obtained using all three techniques together, as well as many of the techniques described in the preceding sections. Because of the building geometry, GNSS ranging (with or without aiding) is typically more accurate in the along-street direction than the across-street direction. Conversely, shadow matching is more accurate in the across-street direction. 3DMA ranging and shadow matching can be integrated simply by forming a weighted average of the two position solutions. The weighting should be directional, which can be achieved either using the covariances of the two position solutions or using the street direction extracted from the mapping. However, where both positioning algorithms score an array of candidate positions, it is better to combine the ranging and shadow matching scores for each candidate and then extract the integrated position solution [100]. Shadow-matching and 3DMA GNSS ranging likelihood surfaces can also be processed with a multi-epoch navigation filter.

Terrain height aiding is inherent in both shadow matching and in 3DMA ranging algorithms that consider multiple candidate positions. By assuming the receive antenna is a fixed distance above the terrain, the grid of candidate positions is constrained to two spatial dimensions, instead of three, reducing the processing load by an order of magnitude. However, a terrain-height-aiding least-squares position solution should also be used to initialize the search region. As it is more accurate, it enables fewer candidates to be considered, again reducing the processing load.

For 3DMA GNSS to be practical, the algorithms must be able to run in real-time on a typical consumer device and have real-time access to suitable 3D mapping data. Computational load is not a major problem for the positioning algorithms. However, ray tracing can be computationally intensive. One solution is to use pre-computed building boundaries, though that can take up more space than the original 3D mapping. Another option is to use projection techniques run on a graphics processing unit (GPU). However, both of these approaches can only predict satellite visibility, not path delays. Real-time path delay determination using ray tracing is currently limited to a hundred or so candidate positions per second.

Highly detailed 3D city models are expensive. However, simple block models, known as level of detail (LOD) 1, are sufficient for most 3D-mapping-aided GNSS implementations. Open Street Map provides freely available building mapping for the world's major cities and many other places, much of it in 3D. Data is also available from national mapping agencies. Although coverage is far from universal, it tends to be available in the dense urban areas where it is most needed. Conventional GNSS positioning usually works well enough in low density areas.

3DMA GNSS can be implemented on either a server or a mobile device. A server-based implementation can use the existing assisted GNSS protocols to communicate with a GNSS receiver, so no modifications to the mobile device would be needed and the 3D mapping would all be kept at the server. However, a server can only provide positioning for a limited number of users at one time, so is best suited to applications requiring only a single-epoch position fix. For continuous navigation and tracking applications, 3D mapping or building boundary data can easily be streamed to users over modern mobile phone connection, assuming an efficient binary format. Pre-loading of data is also possible, but not necessarily convenient as a mobile device could only hold data for a few cities at a time.

8. Summary

Multipath reception is a phenomenon that all GNSS receivers must contend with and for many applications it is the dominate error source. This chapter has provided a survey of the multitude of techniques available to mitigate multipath errors: antenna siting to avoid multipath; antenna types that enhance direct signals and attenuate reflected signals, particularly for fixed sites; adaptive antenna array processing; correlation signal processing; measurement processing techniques like carrier smoothing; navigation processing to de-weight or exclude measurements impacted by multipath; and post-processing and modelling techniques that provide estimates to correct multipath errors. Applicability of these techniques to different GNSS receiver types varies greatly, with mobile phones being especially constrained.

NLOS reception is an additional challenge faced by many applications, especially for users in urban environments. Many receiver multipath mitigation techniques, including antenna and signal processing approaches, do not address NLOS reception. However, navigation processing techniques that help to de-weight or exclude multipath can also be adapted to mitigate NLOS reception. Furthermore, 3DMA techniques, such as shadow matching, are an example of how signal propagation modelling can be applied to actually use NLOS reception and signal blocking as sources of positioning information.

As computational capabilities available to GNSS receiver systems continue to improve, and with increased availability of information from communication networks and aiding sensors, the next decades can expect to see continued improvements in mitigation of multipath and NLOS reception.

9. References

- [1] P. D. Groves, "How Does Non-Line-of-Sight Reception Differ From Multipath Interference?" *Inside GNSS*, Nov/Dec 2013, pp. 40-42, 63.
- [2] B. Rama Rao, W. Kunysz, R. Fante, and K. McDonald, *GPS/GNSS Antennas*, Artech House, 2013.
- [3] K. M. Larson et al., "Environmental Sensing: A Revolution in GNSS Applications," *Inside GNSS*, Jul/Aug 2014, pp. 36-46.

- [4] J. Bradbury, "Prediction of Urban GNSS Availability and Signal Degradation Using Virtual Reality City Models," in *Proc. ION GNSS 2007*, Fort Worth, TX, pp. 2696-2706, September 2007.
- [5] M. S. Braasch, "Multipath Effects," In *Global Positioning System: Theory and Applications Volume I*, B. W. Parkinson and J. J. Spilker, Jr., (eds.), Washington, DC: AIAA, 1996, pp. 547–568.
- [6] P.D. Groves, *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems*, 2nd Edition, Artech House, 2013.
- [7] B. Hannah, Modeling and Simulation of GPS Multipath Propagation, Ph.D. Dissertation, Queensland University of Technology, 2001.
- [8] P. Misra and P. Enge, *Global Positioning System: Signals, Measurements, and Performance*, 2nd Edition, Ganga-Jamuna Press, 2006.
- [9] A. J. Van Dierendonck, P. Fenton, and T. Ford, "Theory and Performance of a Narrow Correlator Spacing in a GPS Receiver," in *NAVIGATION*, Vol. 39, No. 3, pp. 265–283, 1992.
- [10] A. J. Van Dierendonck, "GPS Receivers," in *Global Positioning System: Theory and Applications Volume I*, B. W. Parkinson and J. J. Spilker, Jr., (eds.), Washington, DC: AIAA, 1996, ch. 8, pp. 329–407.
- [11] P. W. Ward, J. W. Betz, and C. J. Hegarty, "Satellite Signal Acquisition, Tracking and Data Demodulation," In *Understanding GPS Principles and Applications*, 2nd ed., E. D. Kaplan and C. J. Hegarty, (eds.), Norwood, MA: Artech House, 2006, ch. 5, pp. 153–241.
- [12] J. W. Betz, *Engineering Satellite-Based Navigation and Timing: Global Navigation Satellite Systems, Signals, and Receivers*, IEEE Press/ Wiley, 2016.
- [13] R. D. J. Van Nee, "GPS Multipath and Satellite Interference," in *Proc. ION 48th Ann. Meeting*, Washington, DC, pp. 167–177, June 1992.
- [14] J. W. Betz, "Binary Offset Carrier Modulation for Radionavigation," *NAVIGATION*, Vol. 48, No. 4, pp. 227–246, 2001.
- [15] M. Irsigler, and B. Eissfeller, "Comparison of Multipath Mitigation Techniques with Consideration of Future Signal Structures," in *Proc. ION GPS/GNSS 2003*, Portland, OR, pp. 2584–2592, September 2003.
- [16] D. B. Thornberg, D. S. Thornberg, M. F. DiBenedetto, M. S. Braasch, F. van Graas, C. Bartone, "LAAS Integrated Multipath-Limiting Antenna", *NAVIGATION*, Journal of The Institute of Navigation, Vol. 50, No. 2, Summer 2003, pp. 117-130.
- [17] F. Dovis, M. Pini and P. Mulassano, "Multiple DLL architecture for multipath recovery in navigation receivers," 2004 IEEE 59th Vehicular Technology Conference. VTC 2004-Spring (IEEE Cat. No.04CH37514), 2004, pp. 2848-2851 Vol.5.
- [18] C. Yang, A. Porter, "GPS Multipath Estimation and Mitigation Via Polarization Sensing Diversity: Parallel Iterative Cross Cancellation," in *Proceedings of the 18th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2005)*, Long Beach, CA, September 2005, pp. 2707-2719.
- [19] Z. Jiang and P. D. Groves, "NLOS GPS signal detection using a dual-polarisation antenna," in *GPS Solutions*, Vol. 18, No. 1, pp. 15–26, 2014 (online: 2012) .
- [20] B. Ashman, Incorporation of GNSS Multipath to Improve Autonomous Rendezvous, Docking and Proximity Operations in Space, Ph.D. Dissertation, Purdue University, 2016.

- [21] S. Katzberg, J. Garrison, and C. Howell, "Simple Over-Water Altimeter Using GPS Reflections," in *Proceedings of the 12th International Technical Meeting of the Satellite Division of the Institute of Navigation*, Nashville, TN, pp. 1819–1828, September 1999.
- [22] D. Masters, P. Axelrad, V. Zavorotny, S. Katzberg and F. Lalezari, "A Passive GPS Bistatic Radar Altimeter for Aircraft Navigation," in *Proceedings of the 14th International Technical Meeting of the Institute of Navigation*. Cambridge, MA: Institute of Navigation, pp. 2435–2445, June 2001.
- [23] E. Vinande, D. Akos, D. Masters, P. Axelrad, and S. Esterhuizen, "GPS Bistatic Radar Measurements of Aircraft Altitude and Ground Objects with a Software Receiver," in *Proceedings of the 61st Annual Meeting of the Institute of Navigation*, Cambridge, MA, pp. 528–534, June 2005.
- [24] J. R. Dowdle and D. E. Gustafson and J. John M. Elwell, "Geographical Navigation Using Multipath Wireless Navigation Signals," US Patent 6,693,592, February 2004.
- [25] I. Cohen, "Relative Navigation for Hubble Servicing Mission Using Reflected GPS Signals," Master's thesis, University of Maryland, 2007.
- [26] I. Cohen and G. Boegner Jr., "Method and Apparatus for Relative Navigation Using Reflected GPS Signals," US Patent 7,817,087, October 2010.
- [27] L. Garin, F. van Diggelen, J-M. Rousseau, "Strobe & Edge Correlator Multipath Mitigation for Code," *Proceedings of the 9th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1996)*, Kansas City, MO, September 1996, pp. 657-664.
- [28] R. R. Hatch, R. G. Keegan, T. A. Stansell, "Leica's Code and Phase Multipath Mitigation Techniques," *Proceedings of the 1997 National Technical Meeting of The Institute of Navigation*, Santa Monica, CA, January 1997, pp. 217-225.
- [29] G. A. McGraw and M. S. Braasch, "GNSS Multipath Mitigation Using Gated and High Resolution Correlator Concepts," in *Proceedings of the Inst. of Navigation National Technical Meeting*, San Diego, CA, Jan. 25-27, 1999.
- [30] G. A. McGraw, "Practical GPS Carrier Phase Multipath Mitigation using High Resolution Correlator Techniques," in *Proceedings of International Association of Institutes of Navigation World Congress / Annual Meeting of the Inst. of Navigation*, San Diego, CA, June 26-28, 2000.
- [31] V. A. Veitsel, A. V. Zhdanov and M. I. Zhodzishsky, "The Mitigation of Multipath Errors by Strobe Correlators in GPS/GLONASS Receivers," *GPS Solutions*, Vol. 2, pp. 38-45, 1998.
- [32] A. V. Zhdanov, V. A. Veitsel, M. I. Zhodzishsky and J. Ashjaee, "Multipath Error Reduction in Signal Processing," *Proceedings of the 12th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1999)*, Nashville, TN, September 1999, pp. 1217-1224.
- [33] M. Irsigler and B. Eissfeller, "Comparison of Multipath Mitigation Techniques with Consideration of Future Signal Structures," *Proceedings of the 16th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS/GNSS 2003)*, Portland, OR, September 2003, pp. 2584-2592.
- [34] P. G. Mattos, P. G., "Multipath elimination for the low-cost consumer GPS," *Proc. ION GPS-96*, Kansas, MO, September 1996, pp. 665–672.
- [35] B. R. Townsend, P. C. Fenton, "A Practical Approach to the Reduction of Pseudorange Multipath Errors in a L1 GPS Receiver," *Proceedings of the 7th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1994)*, Salt Lake City, UT, September 1994, pp. 143-148.

- [36] B. R. Townsend, P. C. Fenton, K. J. Van Dierendonck, R. D. J. van Nee, "Performance Evaluation of the Multipath Estimating Delay Lock Loop", *NAVIGATION, Journal of The Institute of Navigation*, Vol. 42, No. 3, Fall 1995, pp. 503-514.
- [37] L. R. Weill, "Achieving Theoretical Accuracy Limits for Pseudorangeing in the Presence of Multipath," *Proceedings of the 8th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1995)*, Palm Springs, CA, September 1995, pp. 1521-1530.
- [38] L. R. Weill, "Application of Superresolution Concepts to the GPS Multipath Mitigation Problem," *Proceedings of the 1998 National Technical Meeting of The Institute of Navigation*, Long Beach, CA, January 1998, pp. 673-682.
- [39] L. R. Weill, "Multipath Mitigation: How Good Can It Get With New Signals?" *GPS World*, June 2003, pp. 106-113.
- [40] P. C. Fenton, J. Jones, "The Theory and Performance of NovAtel Inc.'s Vision Correlator," *Proceedings of the 18th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2005)*, Long Beach, CA, September 2005, pp. 2178-2186.
- [41] M. Z. H. Bhuiyan, E. S. Lohan, "Multipath Mitigation Techniques for Satellite-Based Positioning Applications," in Jin S (Ed.) *Global Navigation Satellite Systems: Signal, Theory and Applications*, InTech, Rijeka, Croatia, 2012, pp. 405-426
- [42] J. K. Ray, M. E. Cannon, P. C. Fenton, "Mitigation of Static Carrier-Phase Multipath Effects Using Multiple Closely Spaced Antennas", *NAVIGATION, Journal of The Institute of Navigation*, Vol. 46, No. 3, Fall 1999, pp. 193-202.
- [43] K. Gold, A. Brown, "An Array of Digital Antenna Elements for Mitigation of Multipath for Carrier Landings," *Proceedings of the 2005 National Technical Meeting of The Institute of Navigation*, San Diego, CA, January 2005, pp. 190-196.
- [44] McGraw, Gary A., Young, Ryan S.Y., Reichenauer, Kristen, Stevens, Jeff, Ventrone, Fred, "GPS Multipath Mitigation Assessment of Digital Beam Forming Antenna Technology in a JPALS Dual Frequency Smoothing Architecture," *Proceedings of the 2004 National Technical Meeting of The Institute of Navigation*, San Diego, CA, January 2004, pp. 561-572.
- [45] A. Soloviev, and F. van Graas, "Utilizing Multipath Reflections in Deeply Integrated GPS/INS Architecture for Navigation in Urban Environments," *Proc. IEEE/ION PLANS*, Monterey, CA, May 2008, pp. 383-393.
- [46] P. Xie, M. G. Petovello, and C. Basnayake, "Multipath Signal Assessment in the High Sensitivity Receivers for Vehicular Applications," *Proc. ION GNSS 2011*, Portland, OR, pp.1764-1776.
- [47] T. Pany, B. Eissfeller, "Demonstration of a Synthetic Phased Array Antenna for Carrier/Code Multipath Mitigation," *Proc. ION GNSS 2008*, Savannah, GA, September 2008, pp. 663-668.
- [48] S. Draganov, M. Harlacher, L. Haas, "Multipath Mitigation via Synthetic Aperture Beamforming," *Proc. ION GNSS 2009*, Savannah, GA, September 2009, pp. 1707-1715.
- [49] R. Hatch, "The Synergy of GPS Code and Carrier Measurements", in Proc. Of 3rd Int. Symp. on Satellite Doppler Positioning, 1982.
- [50] P. Y. C. Hwang, and R.G. Brown, "GPS Navigation: Combining Pseudorange With Continuous Carrier Phase Using a Kalman Filter", *NAVIGATION, Journal of The Institute of Navigation*, Vol. 37, No. 2, Summer 1990, pp. 181-196.
- [51] R.G. Brown and P.Y.C. Hwang, Introduction to Random Signal Analysis and Applied Kalman Filtering, 3rd Ed., Wiley, 1997.

- [52] P. Y. C. Hwang, G. A. McGraw, and J. R. Bader, "Enhanced Differential GPS Carrier-Smoothed Code Processing Using Dual Frequency Measurements," *NAVIGATION, Journal of The Institute of Navigation*, Vol. 46, No. 2, Summer 1999, pp.127-137.
- [53] G. F. Franklin, J. D. Powell and M. L. Workman, *Digital Control of Dynamic Systems*, 2nd Ed., Addison-Wesley, 1990.
- [54] McGraw, G. A., "Generalized Divergence-Free Carrier Smoothing with Applications to Dual Frequency Differential GPS," *NAVIGATION: Journal of the Institute of Navigation*, Vol. 56, No. 2, Summer 2009, pp. 115-122.
- [55] P. Henkel and C. Günther, "Reliable Integer Ambiguity Resolution: Multi-Frequency Code Carrier Linear Combinations and Statistical A Priori Knowledge of Attitude", *NAVIGATION: Journal of the Institute of Navigation*, Vol. 59, No. 1, Spring 2012, pp. 61-75.
- [56] "Minimum Operational Performance Standards for Global Positioning System/Wide Area Augmentation System Airborne Equipment," RTCA DO-229D, 2006.
- [57] H. Hartinger and F. K. Brunner, "Variances of GPS Phase Observations: The SIGMA- ε Model," in *GPS Solutions*, Vol. 2, No. 3, pp. 35–43, 1999
- [58] P. D. Groves, Z. Jiang, B. Skelton, P. A. Cross, L. Lau, Y. Adane, I. Kale, "Novel Multipath Mitigation Methods using a Dual-Polarization Antenna," in *Proceedings of the 23rd International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2010)*, Portland, OR, September 2010, pp. 140-151.
- [59] N. Viandier et al., "GNSS Performance Enhancement in Urban Environment based on Pseudorange Error Model," *Proc. IEEE/ION PLANS*, Monterey, CA, pp. 377–382, May 2008.
- [60] P. R. R. Strode and P. D. Groves, "GNSS multipath detection using three-frequency signal-to-noise measurements," in *GPS Solutions*, Vol. 20, No. 3, pp. 399–412, 2016 (online: 2015).
- [61] S. Hilla and M. Cline, "Evaluating pseudorange multipath effects at stations in the National CORS network," in *GPS Solutions*, Vol. 7, No. 4, pp. 253–267, 2004.
- [62] Z. Jiang, P. D. Groves, W. Y. Ochieng, S. Feng, C. D. Milner, and P. G. Mattos, "Multi-Constellation GNSS Multipath Mitigation Using Consistency Checking," *Proc. ION GNSS 2011*, Portland, OR, pp. 3889-3902, September 2011.
- [63] P. D. Groves and Z. Jiang, "Height Aiding, C/N0 Weighting and Consistency Checking for GNSS NLOS and Multipath Mitigation in Urban Areas," in *Journal of Navigation*, Vol. 66, No. 5, pp. 653-669, 2013.
- [64] P. H. S. Torr, and A. Zisserman, "MLE-SAC: A New Robust Estimator with Application to Estimating Image Geometry," in *Computer Vision and Image Understanding*, Vol. 78, No. 1, pp 138-156, 2000.
- [65] L-T. Hsu, H Tokura, N. Kubo, Y, Gu, and S. Kmaijo, "Multiple Faulty GNSS Measurement Exclusion Based on Consistency Check in Urban Canyons" in *IEEE Sensors Journal*, Vol. 17. No. 6, pp. 1909-1917, 2017.
- [66] M. Spangenberg, et al., "Detection of Variance Changes and Mean Value Jumps in Measurement Noise for Multipath Mitigation in Urban Navigation," in *NAVIGATION*, Vol. 57, No. 1, pp. 35-52, 2010.
- [67] M. Lashley and D. M. Bevly, "Comparison in the Performance of the Vector Delay/Frequency Lock Loop and Equivalent Scalar Tracking Loops in Dense Foliage and Urban Canyon," *Proc. ION GNSS 2011*, Portland, OR, pp. 1786-1803, September 2011.

- [68] L.-T. Hsu, S.-S. Jan, P. D. Groves, and N. Kubo, "Multipath mitigation and NLOS detection using vector tracking in urban environments," in *GPS Solutions*, Vol. 19, No. 2, pp. 249–262, 2015 (online: 2014).
- [69] J. A. Farrell, *Aided Navigation: GPS with High Rate Sensors*, McGraw Hill, 2008.
- [70] M. S. Grewal, A. P. Andrews, and C. G. Bartone, *Global Positioning Systems, Inertial Navigation, and Integration*, 3rd ed., Wiley, 2013.
- [71] J. Marais, M. Berbineau, and M. Heddebaut, "Land Mobile GNSS Availability and Multipath Evaluation Tool," in *IEEE Transactions on Vehicular Technology*, Vol. 54, No. 5, pp. 1697-1704, 2005.
- [72] J. Meguro, et al., "GPS Multipath Mitigation for Urban Area Using Omnidirectional Infrared Camera," in *IEEE Transactions on Intelligent Transportation Systems*, Vol. 10, No. 1, pp. 22-30, 2009.
- [73] T.A. Springer, "High Accuracy GNSS Solutions and Services", *Global Navigation Satellite Systems Overview, GNSS Overview*, 6 Feb. 2017, www.positim.com/navsys_overview.html.
- [74] S. Jan, and T. An-Lin. "Comprehensive Comparisons of Satellite Data, Signals, and Measurements between the BeiDou Navigation Satellite System and the Global Positioning System", *Sensors* (Basel, Switzerland), MDPI, 13 May 2016, www.ncbi.nlm.nih.gov/pmc/articles/PMC4883380/.
- [75] J. F. Genrich, and Y. Bock, Rapid Resolution of Crustal Motion at Short Ranges with the Global Positioning System, *Journal of Geophysical Research*, Vol. 97, pp. 3261-3269, 1992.
- [76] C. Atkins, *Observation-Domain Sidereal Filtering for High-Rate GPS Precise Point Positioning*, Ph.D. Dissertation, University College London, 2016.
- [77] C. Larson, A. Bilich, and P. Axelrad, "Improving the Precision of High-rate GPS," *Journal of Geophysical Research*, Vol. 112, No. B05422, pp. 1-11, 2007.
- [78] P. Axelrad, K. M. Larson, Kristine, B. Jones, "Use of the Correct Satellite Repeat Period to Characterize and Reduce Site-Specific Multipath Errors," *Proceedings of the 18th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2005)*, Long Beach, CA, pp. 2638-2648, September 2005.
- [79] K. Choi, A. Bilich, K. M. Larson, and P. Axelrad, Modified Sidereal Filtering: Implications for High-Rate GPS Positioning, *Geophysical Research Letters*, 31, L22608, 2004.
- [80] D. C. Agnew, and K. M. Larson, Finding the Repeat Times of the GPS Constellation, *GPS Solutions*, 11: 71-76, 2007.
- [81] G. J. Bishop, D. S. Coco, P. H. Kappler, and E. A. Holland, "Studies and Performance of a New Technique for Mitigation of Pseudorange Multipath Effects in GPS Ground Stations," *Proceedings of the 1994 National Technical Meeting of the Institute of Navigation*, San Diego, CA, pp. 231-242, January 1994.
- [82] A. K. Reichert, and P. Axelrad, "Carrier-Phase Multipath Corrections for GPS-based Satellite Attitude Determination," *Navigation: JION*, Vol. 48, No. 2, 2001, pp. 77–88.
- [83] S. Hodgart, and R. Wong, "Statistically Optimized In-Flight Estimation of GPS Carrier Phase Multipath for LEO Satellite Attitude Determination," *Navigation: JION*, Vol. 53, No. 3, 2006, pp. 181–202.
- [84] M. Psiaki, T. Ertan, B. O'Hanlon, and S. Powell, GNSS Multipath Mitigation using Antenna Motion, *Journal of the Institute of Navigation*, Vol. 62, No. 1, pp. 1–22, Spring 2015.
- [85] J. Goodman, *Introduction to Fourier Optics*, Englewood, CO: Roberts & Co., 2005.

- [86] J. Keller, Geometrical Theory of Diffraction, *Journal of the Optical Society of America*, Vol. 52, No. 2, pp. 116-130, 1961.
- [87] V. A. Borovikov and B. Ye. Kinber, *Geometrical Theory of Diffraction*. London: The Institute of Electrical Engineers, 1994.
- [88] S. Gomez, "Three Years of Global Positioning Systems Experience on International Space Station", NASA Johnson Space Center, NASA Technical Publication NASA/TP-2006-213168, 2006.
- [89] S. Byun, and G. Hajj, and L. Young, "Development and Application of GPS Signal Multipath Simulator", *Radio Science*, Vol. 37, No. 6, pp. 10.1-10.23, November 2002.
- [90] J. Weiss, Modeling and Characterization of Multipath in Global Navigation Satellite System Ranging Signals, Ph.D. dissertation, University of Colorado Boulder, 2007.
- [91] L. Lau, and P. Cross, Development and Testing of a New Ray-Tracing Approach to GNSS Carrier-Phase Multipath Modeling, *Journal of Geodesy*, 81(11), pp. 713-732, 2007.
- [92] S. Gomez, and S. Hwu, "Comparison of Space Shuttle GPS Flight Data to Geometric Theory of Diffraction Predictions", in *Proceedings of the 10th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1997)*, Kansas City, MO, September 1997.
- [93] P. Axelrad, K. Gold, P. Madhani, and A. Reichert, "Analysis of Orbit Errors Induced by Multipath for the ICESat Observatory," *Proceedings of the 12th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1999)*, Nashville, TN, pp. 875-884, September 1999.
- [94] M. Braasch, Isolation of GPS Multipath and Receiver Tracking Errors, *Navigation, Journal of the Institute of Navigation*, Vol. 41, No. 4, pp. 415-434, Winter 1994-1995.
- [95] M. Braasch, "Multipath", *Springer Handbook of Global Navigation Satellite Systems*, ed. P. Teunissen, and O. Montenbruck, Springer, pp. 445-470, 2017.
- [96] S. Bisnath and R. Langley, "Pseudorange Multipath Mitigation By Means of Multipath Monitoring and De-Weighting," in *Proceedings of the International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation*, Banff, Alberta, pp. 392-400, June 2001.
- [97] C. Comp, "GPS Carrier Phase Multipath Characterization and a Mitigation Technique Using the Signal-to-Noise Ratio," Ph.D. dissertation, University of Colorado, 1996.
- [98] P. Axelrad, C. J. Comp and P. F. Macdoran, "SNR-based multipath error correction for GPS differential phase," in *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 32, No. 2, pp. 650-660, April 1996.
- [99] A. Simsky, "Three's the Charm: Triple-Frequency Combinations in Future GNSS," *Inside GNSS*, July/August 2006, pp. 38-41.
- [100] M. Adjrak and P. D. Groves, "Intelligent Urban Positioning using Shadow Matching and GNSS Ranging Aided by 3D Mapping," Proc. ION GNSS+ 2016, Portland, OR, September 2016.
- [101] M. Adjrak and P. D. Groves, "Enhancing Least Squares GNSS Positioning with 3D Mapping without Accurate Prior Knowledge," in *NAVIGATION*, Vol. 64, No. 1, pp. 75-91, 2017.
- [102] J. R. Amt, and J. F. Raquet, "Positioning for Range-Based Land Navigation Systems Using Surface Topography," Proc. ION GNSS 2006, Fort Worth, TX, pp. 1494-1505, September 2006.
- [103] T. Iwase, N. Suzuki, and Y. Watanabe, "Estimation and exclusion of multipath range error for robust positioning," in *GPS Solutions*, Vol. 17, No. 1, pp. 53-62, 2013 (online: 2012).
- [104] M. Obst, S. Bauer, and G. Wanielik, "Urban Multipath Detection and mitigation with Dynamic 3D Maps for Reliable Land Vehicle Localization," Proc. IEEE/ION PLANS 2012, Myrtle Beach, SC, pp.685-691, April 2012.

- [105] S., Peyraud, et al., “About Non-Line-Of-Sight Satellite Detection and Exclusion in a 3D Map-Aided Localization Algorithm,” in *Sensors*, Vol. 13, pp. 829-847, 2013.
- [106] P. D. Groves, and M. Adjrak, “Likelihood-based GNSS positioning using LOS/NLOS predictions from 3D mapping and pseudoranges” in *GPS Solutions*, Vol. 21, No. 4, pp. 1805-1816, 2017.
- [107] T. Suzuki, and N. Kubo, “Correcting GNSS Multipath Errors Using a 3D Surface Model and Particle Filter,” *Proc. ION GNSS+ 2013*, Nashville, TN, pp. 1583–1595, September 2013.
- [108] R. Kumar and M. G., Petovello, “A Novel GNSS Positioning Technique for Improved Accuracy in Urban Canyon Scenarios using 3D City Model”, *Proc. ION GNSS+ 2014*, Tampa, FL, pp. 2139–2148, September 2014.
- [109] L.-T. Hsu, Y. Gu and S. Kamijo, “3D building model-based pedestrian positioning method using GPS/GLOANSS/QZSS and its reliability calculation”, in *GPS Solutions*, Vol. 20, No. 3, pp. 413–428, 2016 (online: 2015).
- [110] P. D. Groves, “Shadow Matching: A New GNSS Positioning Technique for Urban Canyons”. in *The Journal of Navigation*, Vol. 64, No. 3, pp. 417-430, 2011.
- [111] P. D. Groves, L. Wang, M. Adjrak and C. Ellul, “GNSS Shadow Matching: The Challenges Ahead,” *Proc. ION GNSS+ 2015*, Tampa, FL, pp. 2421–2443, September 2015.
- [112] L. Wang, P. D. Groves, and M. K. Ziebart, “Smartphone Shadow Matching for Better Cross-street GNSS Positioning in Urban Environments”. in *The Journal of Navigation*, Vol. 68, No. 3, pp. 411–433, 2015.
- [113] H. Gao and P. D. Groves, “Environmental Context Detection for Adaptive Navigation using GNSS Measurements from a Smartphone,” *NAVIGATION*, Vol. 65, No. 1, pp. 99–116, 2018.
- [114] T. Suzuki, and N. Kubo, “GNSS Positioning with Multipath Simulation using 3D Surface Model in Urban Canyon,” *Proc. ION GNSS 2012*, Nashville, TN, pp. 438–447, September 2012.
- [115] J. T. Isaacs, A. T. Irish, et al., “Bayesian localization and mapping using GNSS SNR measurements”. *Proc. IEEE/ION PLANS 2014*. Monterey, CA, pp. 445–451, May 2014.
- [116] L. Wang, Investigation of Shadow Matching for GNSS Positioning in Urban Canyons, PhD Thesis, University College London, <http://discovery.ucl.ac.uk>, 2015.
- [117] R. Yozevitch, and B. Ben-Moshe, “A Robust Shadow Matching Algorithm for GNSS Positioning,” in *NAVIGATION*, Vol. 62, No. 2, pp. 95–109, 2015.